Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE

Lenka Slámová

MODELOVÁNÍ ÚROKOVÝCH SAZEB S VYUŽITÍM LÉVYHO PROCESŮ

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce:	Prof. RNDr. Viktor Beneš, DrSc.
Studijní program:	Matematika
Studijní obor:	Pravděpodobnost, matematická statistika a ekonometrie
Studijní plán:	Ekonometrie

2010

Charles University in Prague
Faculty of Mathematics and Physics

DIPLOMA THESIS

Lenka Slámová

INTEREST RATE MODELLING WITH LÉVY PROCESSES

Department of Probability and Mathematical Statistics

Thesis supervisor:	Prof. RNDr. Viktor Beneš, DrSc.
Study programme:	Mathematics
Study branch:	Probability, mathematical statistics and econometrics
Study plan:	Econometrics

2010

Poděkování

Na tomto místě bych ráda poděkovala svému vedoucímu prof. RNDr. Viktoru Benešovi, DrSc., za ochotu při vedení práce, svému školiteli Dr. Michaelovi Schröderovi z VU University Amsterdam za velmi zajímavé téma a nápady pro další rozvoj práce. Dále bych chtěla poděkovat Lukáši Poulovi za veškerou pomoc při psaní práce.

Největší dík patří mé rodině, mé mamince, tatínkovi a sestřičce, za morální podporu během studií.

Prohlášení

Prohlašuji, že jsem svou diplomovou práci napsala samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

V Praze dne 12.4.2010 Lenka Slámová

Contents

Název práce: Modelování úrokových měr s využitím Lévyho procesů
Autor: Lenka Slámová
Katedra: Katedra pravděpodobnosti a matematické statistiky
Vedoucí diplomové práce: Prof. RNDr. Viktor Beneš, DrSc.
E-mail vedoucího: benesv@karlin.mff.cuni.cz

Abstrakt: V předložené práci studujeme HJM model časové struktury úrokových sazeb řízený Lévyho procesem. Studujeme bezaritrážní dynamiku diskontovaných cen bezkupónových dluhopisů a jako důsledek obdržíme model pro proces bezrizikové úrokové sazby. Speciálně se zaměříme na proces krátkodobé úrokové sazby a zformulujeme kritéria pro tzv. mean reversion. Teorie nám dává postup pro získání procesu krátkodobé úrokové sazby pro obecný Lévyho řídící proces a obecnou strukturu volatility, a neprázdnost této teorie demonstrujeme na příkladu Orstein–Uhlenbeckova procesu řízeného Lévyho procesem, s marginálním generalized inverse Gaussian rozdělením. Výsledkem je explicitní vzorec pro proces krátkodobé úrokové sazby, který zobecňuje Vašíčkův model, a navíc je vždy kladný. Nakonec studujeme numerické metody pro takto zkonstruovaný proces úrokových sazeb, jako jsou simulace a multinomické stromy.

Klíčová slova: Časová struktura úrokových sazeb, short rate proces, Lévyho procesy, procesy Ornstein–Uhlenbeckova typu, generalized inverse Gaussian rozdělení

Title: Interest rate modelling with Lévy processes
Author: Lenka Slámová
Department: Department of probability and mathematical statistics
Supervisor: Prof. RNDr. Viktor Beneš, DrSc.
Supervisor's e-mail adress: benesv@karlin.mff.cuni.cz

Abstract: In this work we study the HJM model of the term structure of interest rates driven by a Lévy process. We study the no-arbitrage dynamics of the discounted bond prices and obtain a risk-neutral dynamics of the short rate as a consequence. We study in particular the short rate process and formulate a criteria for mean reversion. The theory gives us a machinery producing short rate processes associated with a general Lévy driver and general volatility structure and we show the non-emptiness of the theory by demonstrating the previous on an example of an OU type process associated with the generalized inverse Gaussian distribution. The upshot is an explicitely given short rate process that generalizes the Vašíček model, and moreover stays positive. Finally we study numerical methods for thus constructed short rate process such as simulations and lattice approximations.

Keywords: Term structure of interest rates, short rate process, Lévy processes, Ornstein-Uhlenbeck type processes, generalized inverse Gaussian distribution

Notation

$(\mathbb{P}, \Omega, \mathbb{F}, \mathcal{F})$	filtered probability space
E	$\mathbb{P}$-expected value
φ	density function of the Gaussian $\mathrm{N}(0,1)$ distribution
Φ	cumulative distribution function of the Gaussian $\mathrm{N}(0,1)$ distribution
$\widehat{\mu}$	characteristic function of a probability measure μ
$\widehat{P}$	characteristic function of the distribution P
Ψ_X	characteristic function of a random variable X
Θ_X	logarithm of the characteristic function of a random variable X
ψ_X	moment generating function of a random variable X
θ_X	logarithm of moment generating function of a random variable X
$\mathbb{R}$	set of real numbers $(-\infty, \infty)$
$\mathbb{R}^+$	set of positive real numbers $(0, \infty)$
$\mathbb{C}$	set of complex numbers
$\mathbb{N}$	set of positive integers $\{1, 2, 3, \dots\}$
$L^p(A)$	space of p-integrable functions $f : A \to \mathbb{R}$
$L^p_{\mathrm{bd}}(A)$	space of p-integrable and bounded functions $f : A \to \mathbb{R}$
$\mathcal{C}^k(A)$	space of k-times continuously differentiable functions on $A \subset \mathbb{R}$
Dom	Domain of a function
Im	imaginary part
Re	real part
$\mathrm{N}(a, b)$	Gaussian distribution with mean a and variance b
$\mathrm{Po}(\lambda)$	Poisson distribution with intensity λ
GIG, IG, Γ	generalized inverse Gaussian, inverse Gaussian and Gamma distribution
NIG	normal inverse Gaussian distribution
$P(t, T)$	time t price of a zero coupon bond maturing at time T
B_t	time t value of a savings account
$f(t, T)$	(instantaneous) forward rate at time t for period $[T, T + \mathrm{d}T)$
r_t	(instantaneous) short rate at time t for period $[t, t + \mathrm{d}t)$
RHS	right hand side of an equation
LHS	left hand side of an equation

Chapter 1

Introduction

In the last 40 years a lot of models describing the dynamics of interest rates have been developed. A very popular and widely accepted class of interest rate models are those where we describe the short rate dynamics via a one dimensional diffusion process depending on some parameters; we call such models one-factor short rate models. The pioneer approaches were proposed by Vašíček (1977), Dothan (1978) and Cox et al. (1985); Hull and White (1990) generalized and extended the classical Vašíček and Cox, Ingresoll and Ross models. All these models are driven by a Brownian motion, a notion well understood and widely used in practice. However, it is known for a long time that the Gaussian distribution of the driving process does not fit very well the real financial data. A natural step was to find a different driving process. Starting with variance gamma process in Madan and Senete (1990) for describing the stock prices, stochastic volatility model of the stock price driven by normal inverse Gaussian process was introduced in Barndorff-Nielsen (1998). Eberlein and Raible (1999) came up with a term structure model driven by generalized hyperbolic processes that was developed in a serie of articles.

The present thesis, partly inspired by the work of Eberlein and Raible (1999), is concerned with modelling of short rates which arise from a bond market equilibrium, and develops a Lévy process perspective on it. Concentrating on a class of Lévy driven "affine models" the first idea is as follows. We assume that bond prices are determined by forward rates and model forward rates as a Lévy driven affine process. By choice of a specific form of the drift term of this process we achieve arbitrage-free bond prices. As a consequence we obtain a short rate process whose dynamics by construction originates from an arbitrage-free bond market. From an empirical point of view we request short rates to satisfy "stylized facts" such as positivity and convergence in the long run to a "historical mean rate". The second idea is to effect this analysis in terms of the characteristic function of the Lévy process driving the dynamics. A principal problem is to provide concrete examples of short rate processes which arise from the above no-arbitrage framework and satisfy the above stylized facts as well. We address non-emptiness of our approach by providing concrete examples of Lévy driven short rates resulting from the analysis, in particular by asking for analogues to the Vašíček short rate models in the Lévy context. This suggest a programme for construction arbitrage-free short rate models.

The main result of the thesis is a realisation of this programme that is developed in Chapter 5. More precisely, in Section 5.2, as a first step of the programme, we work with affine Lévy driven process that are the analogues to Itô processes with deterministic coefficients, and we obtain in fact a "short rate machine". We are able to construct a short rate process originating from a bond market equilibrium and associated with an arbitrary Lévy process and with an arbitrary volatility structure. As a second step of the programme, we develop in Section 5.3 criteria for mean convergence of the short rate processes constructed in step 1 in terms of the moment generating function of the Lévy drivers of their dynamics. Non-emptiness of our approach, step three of our programme, is addressed in Section 5.4. Here we use our construction to obtain short rate processes of Ornstein–Uhlenbeck type associated with so called generalized inverse Gaussian (GIG) and normal inverse Gaussian (NIG) distributions. This short rate models provide analogues to the Vašíček model of the short rate and in the GIG case avoid the deficiencies this model is known for: we have now positive short rate processes that converge to positive mean rates.

Background material about Lévy processes, mostly collected from Sato (1999) and Bertoin (1996), in particular subordinators such as generalized inverse Gaussian processes, is collected in Chapter 2. In Chapter 3 we give a brief review of the stochastic integration theory with respect to semimartingales which is essential to constructing the class of Lévy driven process we consider. Here we follow Jacod and Shiryaev (2003). The contribution to the theory is a detailed proof of so called Key theorem from Eberlein and Raible (1999) and a proof of independent increments property of Lévy driven stochastic integrals. The material about Ornstein–Uhlenbeck type processes is summarized in Chapter 4. This chapter contains results about OU type processes introduced by Barndorff-Nielsen et al. (1998) and about relationships between the OU type process and the driving Lévy process. A study and a summary of properties of OU type processes associated with generalized inverse Gaussian distributions is a new contribution. Practical issues of the models developed in Chapter 5 are studied in Chapters 6 and 7. Chapter 6 solves the problem of calibration of the short rate models using the market yield curves. Following the work Raible (2000) and Eberlein and Kluge (2006) we show how to estimate the parameters of one of the models studied in Chapter 5. The maximum likelihood method is used to calibrate the OU–NIG model of the short rate based on the real market data. Numerical methods for practical use of the models, such as derivatives pricing, are develloped in Chapter 7. We give algorithms for Monte Carlo simulations of the Lévy driven short rate processes from Chapter 5 and also provide, as a last set of own results, a construction of multinomial tree approximations for the Lévy driven short rate processes.

Chapter 2

Lévy processes and their principal structure

In this chapter we review pertinent notions and results about Lévy processes. We use as our main references Sato (1999), Bertoin (1996); but see also Applebaum (2004) and Schoutens (2003).

2.1 Lévy processes and infinite divisibility

Lévy process Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equiped with a filtration $(\mathcal{F}_t, t \geq 0)$ that satisfies the usual conditions (i.e., $\mathcal{F}_0$ contains all null sets and $(\mathcal{F}_t, t \geq 0)$ is right continuous.).

Definition 2.1. We call a stochastic process $(X_t : t \geq 0)$ on $\mathbb{R}$ a *Lévy process* if it satisfies the following conditions:

(1) For any choice of $n \geq 1$ and $0 \leq t_0 < t_1 < \cdots < t_n$, random variables

$$X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}}$$

are independent.

(2) $X_0 = 0$ a.s.

(3) The distribution of $X_{s+t} - X_s$ does not depend on s.

(4) X is stochastically continuous, i.e. for every $t \geq 0$ and $\varepsilon > 0$,

$$\lim_{s \to t} \mathbb{P}[|X_s - X_t| > \varepsilon] = 0.$$

(5) There is $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}[\Omega_0] = 1$ such that, for every $\omega \in \Omega_0$, $X_t(\omega)$ is right-continuous in $t \geq 0$ and has left limits in $t > 0$.

The condition (1) is called *independent increments property* and the condition (3) is called *stationary increments property*. A process having the properties described in condition (5) is called a *càdlàg* process (continue à droite, limite à gauche - right continuous, limits from the left).

Characteristic function In the sequel we will need the notion of a characteristic function, a characteristic exponent and a moment generating function of a random variable, all together called cumulant transforms. The *characteristic function* $\hat{\mu}(z)$ of a probability measure μ on $\mathbb{R}$ is

$$\hat{\mu}(u) = \int_{\mathbb{R}} e^{\mathrm{i}ux} \mu(\mathrm{d}x), \qquad u \in \mathbb{R}.$$

The characteristic function of a random variable X on $\mathbb{R}$ is denoted by Ψ_X and defined by

$$\Psi_X(u) = \int_{\mathbb{R}} e^{\mathrm{i}ux} P_X(\mathrm{d}x) = \mathrm{E}[e^{\mathrm{i}uX}], \qquad u \in \mathbb{R}.$$

We denote the logarithm of the characteristic function of a random variable X by Θ_X and we call it the *characteristic exponent*. We have

$$\Theta_X(u) = \log \mathrm{E}[e^{\mathrm{i}uX}], \qquad u \in \mathbb{R}.$$

Moment generating function The *moment generating function* of a random variable X on $\mathbb{R}$ is denoted by ψ_X and is defined as

$$\psi_X(u) = \mathrm{E}\left[e^{uX}\right], \qquad u \in \mathbb{R}.$$

The logarithm of the moment generating function of a random variable X is denoted by θ_X, and we call it the *log–moment generating function*. We have

$$\theta_X(u) = \log \mathrm{E}\left[e^{uX}\right], \qquad u \in \mathbb{R}.$$

If we extend the characteristic and moment generating functions to a complex plane, one has the following relationships:

$$\Psi_X(u) = \psi_X(\mathrm{i}u) \quad \text{and} \quad \psi_X(u) = \Psi_X(-\mathrm{i}u) \qquad u \in \mathbb{R}.$$

Infinitely divisible distributions There is a one to one correspondence between Lévy processes and infinitely divisible distributions.

Definition 2.2. We say that a random variable X has *infinitely divisible distribution* if, for every $n = 1, 2, \cdots$, there exists a sequence of independent and identically distributed random variables $X_{1,n}, X_{2,n}, \cdots, X_{n,n}$ such that

$$X \stackrel{d}{=} X_{1,n} + X_{2,n} + \cdots + X_{n,n}.$$

We can also say that a probability measure μ on $\mathbb{R}$ is infinitely divisible, if for every $n = 1, 2, \cdots$, there is a probability measure μ_n on $\mathbb{R}$ such that $\mu = \mu_n^{n*}$, where μ^{n*} means the n-th convolution. It is easy to prove that if μ is infinitely divisible, then, for every $t \in [0, \infty)$, μ^{t*} is definable and also infinitely divisible.

Lemma 2.3 (Sato, 1999, Theorem 7.10)**.** *Let (L_t) be a Lévy process. Then for every t, L_t has an infinitely divisible distribution. Conversely, if μ is an infinitely divisible distribution then there exists a Lévy process (L_t) such that the distribution of L_1 is given by μ.*

Sketch of proof. Consider an arbitrary Lévy process $(L_t : t \geq 0)$. Using the decomposition

$$L_1 = L_{1/n} + (L_{2/n} - L_{1/n}) + \cdots + (L_{n/n} - L_{(n-1)/n}),$$

we observe that the disrtibution of L_1 is infinitely divisible. By a similar argument we can see that for any rational number $t \geq 0$, L_t is infinitely divisible as well and its characteristic function is given by

$$\Psi_{L_t}(u) = [\Psi_{L_1}(u)]^t \,. \tag{2.1}$$

Because L is right continuous a.s., the mapping $t \mapsto \Psi_{L_t}(u)$ is right continuous and (2.1) holds for all $t \geq 0$. On the other side consider an arbitrary infinitely divisible distribution μ on $\mathbb{R}$. We may define a Lévy process L_t such that $P_{L_1} = \mu$. □

2.2 Lévy-Khintchine formula and Lévy-Itô decomposition

Lévy-Khintchine formula The following *Lévy-Khintchine representation* gives a representation of characteristic functions of all infinitely divisible distributions, and hence of all Lévy processes. The proof can be found for example in Sato (1999, Theorem 8.1). Denote $D = \{x \in \mathbb{R} : |x| \leq 1\}$.

Theorem 2.4 (Lévy-Khintchine formula)**.** *If μ is an infinitely divisible distribution on $\mathbb{R}$, then*

$$\hat{\mu}(z) = \exp\left[\mathrm{i}az - \frac{1}{2}\sigma z^2 + \int_{\mathbb{R}} (e^{\mathrm{i}zx} - 1 - \mathrm{i}zx1_D(x))U(\mathrm{d}x)\right], \qquad z \in \mathbb{R}, \tag{2.2}$$

where $a \in \mathbb{R}$, $\sigma \geq 0$ and U is a measure on $\mathbb{R}$ satisfying

$$U(\{0\}) = 0 \;\; \textit{and} \;\; \int_{\mathbb{R}} (|x|^2 \wedge 1)U(\mathrm{d}x) < \infty. \tag{2.3}$$

Moreover, the representation of $\hat{\mu}(z)$ by a, σ and U is unique. Conversely, if $\sigma \geq 0$, U is a measure satisfying (2.3) *and $a \in \mathbb{R}$, then there exists an infinitely divisible distribution μ whose characteristic function is given by* (2.2)*.*

The triplet (a, σ, U) in Theorem 2.4 is called the *generating triplet* of μ, σ is called the *Gaussian variance* and the measure U is called the *Lévy measure* of μ. Every Lévy process $(L_t : t \geq 0)$ can be characterised by its generating triplet (a, σ, U) corresponding to L_1, the generating triplet of L_t is then $(ta, t\sigma, tU)$.

If U satisfies an additional condition $\int_{|x|\leq 1} |x| U(\mathrm{d}x) < \infty$, then we get

$$\hat{\mu}(z) = \exp\left[\mathrm{i}a_0 z - \frac{1}{2}\sigma z^2 + \int_{\mathbb{R}} (e^{\mathrm{i}zx} - 1)U(\mathrm{d}x)\right], \qquad z \in \mathbb{R},$$

and the constant a_0 is called the *drift* of μ.

If U satisfies $\int_{|x|>1} |x| U(\mathrm{d}x) < \infty$, then we get

$$\hat{\mu}(z) = \exp\left[\mathrm{i}a_1 z - \frac{1}{2}\sigma z^2 + \int_{\mathbb{R}} (e^{\mathrm{i}zx} - 1 - \mathrm{i}zx)U(\mathrm{d}x)\right], \qquad z \in \mathbb{R},$$

and the constant a_1 is called the *center* of μ. (Sato, 1999, Example 25.12) showed that $a_1 = \int_{\mathbb{R}} x\mu(\mathrm{d}x)$, i.e., the center and the mean of the distribution μ are identical.

If the Lévy measure is of the form $U(\mathrm{d}x) = u(x)\mathrm{d}x$ we call $u(x)$ the *Lévy density.* The Lévy density has the same mathematical requirements as a probability density, except that it does not need to be integrable and must have zero mass at the origin.

Remark 2.5. We can rewrite the characteristic function $\hat{\mu}$ from (2.2) in the following form or decomposition

$$\hat{\mu}(z) = \Psi_1(z)\ \Psi_2(z)\ \Psi_3(z),$$

where

$$\log \Psi_1(z) = \mathrm{i}az - \frac{1}{2}\sigma z^2, \tag{2.4}$$

$$\log \Psi_2(z) = \int_{|x|>1} \left(e^{\mathrm{i}zx} - 1\right) U(\mathrm{d}x), \tag{2.5}$$

$$\log \Psi_3(z) = \int_{|x|\le 1} \left(e^{\mathrm{i}zx} - 1 - \mathrm{i}zx\right) U(\mathrm{d}x). \tag{2.6}$$

Jumps of a Lévy process Let $E \subset \mathbb{R}$ and μ be a given (positive) Radon measure on $(E, \mathcal{E})$. A family of integer-valued random variables $\{N(B), B \in \mathcal{F}\}$ is called a *Poisson random measure* on Ω with *intensity measure* μ, if the following hold:

(1) for every B, $N(B)$ has a Poisson distribution with mean $\mu(B)$;

(2) if $B_1, \cdots, B_n$ are disjoint, then $N(B_1), \cdots, N(B_n)$ are independent;

(3) for every ω, $N(\cdot, \omega)$ is a measure on Ω.

Consider a Lévy process $(L_t : t \ge 0)$ and denote its jumps as

$$\Delta_t(\omega) = L_t(\omega) - L_{t-}(\omega), \qquad \omega \in \Omega_1.$$

For any measurable set $B \subset [0, \infty) \times \mathbb{R}$ we define the *jump measure* of (L_t) as

$$J(B, \omega) = \begin{cases} \#\left\{s : (s, \Delta_s(\omega)) \in B\right\} & \text{for } \omega \in \Omega_1 \\ 0 & \text{for } \omega \notin \Omega_1, \end{cases} \tag{2.7}$$

using Ω_1 from the definition of a Lévy process 2.1. For every measurable set $A \subset \mathbb{R}$, $J([t_1, t_2] \times A)$ counts the number of jumps of L_t between times t_1 and t_2 such that their sizes are in A.

The Lévy measure U of the Lévy process (L_t) (i.e. the Lévy measure of the distribution of L_1) satisfies

$$U(A) = \mathrm{E}\big[\#\{t \in [0,1] : \Delta_t \ne 0, \Delta_t \in A\}\big], \qquad A \in \mathcal{B}(\mathbb{R}).$$

Lévy-Itô decomposition While proving the Lévy-Khintchine formula, another insight into the structure of Lévy processes can be obtained. Every Lévy process can be uniquely decomposed as a sum of three independent process, as shown in the following result. The proof can be found in (Bertoin, 1996, Chapter I, Theorem 1).

Theorem 2.6 (Lévy-Itô decompositon). *Let $(L_t : t \geq 0)$ be a Lévy process on $\mathbb{R}$ with a generating triplet (a, σ, U). Then*

(1) The jump measure J of (L_t) is a Poisson random measure on $[0, \infty) \times \mathbb{R}$ with intensity measure $U(\mathrm{d}x)\mathrm{d}t$.

(2) We have

$$L_t = L_t^{(1)} + L_t^{(2)} + L_t^{(3)}, \tag{2.8}$$

where

$$\begin{aligned}
L_t^{(1)} &= at + \sqrt{\sigma} W_t \\
L_t^{(2)} &= \sum_{s \leq t} \Delta_s \mathbf{1}_{\{|\Delta_s| \geq 1\}} = \int_{s \leq t, |x| \geq 1} x J(\mathrm{d}s \times \mathrm{d}x) \\
L_t^{(3)} &= \lim_{\varepsilon \downarrow 0} \left(\sum_{s \leq t} \Delta_s \mathbf{1}_{\{\varepsilon < |\Delta_s| < 1\}} - t \int_{\varepsilon < |x| < 1} x U(\mathrm{d}x) \right) \\
&= \lim_{\varepsilon \downarrow 0} \int_{s \leq t,\ \varepsilon < |x| < 1} x \{ J(\mathrm{d}s \times \mathrm{d}x) - U(\mathrm{d}x)\mathrm{d}s \}.
\end{aligned}$$

The convergence of the third component is almost sure and uniform in t.

The first component is the continuous part of (L_t), $L_t^{(1)}$ is a linear transform of a Brownian motion with drift and its characteristic function is given by Ψ_1 in (2.4). The second and third components are the jump parts, $L_t^{(2)}$ is a compound Poisson process incorporating only jumps of size at least 1 and its characteristic function is Ψ_2 in (2.5). Finally $L_t^{(3)}$ is a pure jump martingale having only jumps of size less than 1 with characteristic function Ψ_3 in (2.6). We call the process of type $L_t^{(3)}$ a compensated compound Poisson process.

2.3 Classification of Lévy processes

Lévy processes can be classified into several classes based on their probabilistic properties. According to Sato (1999) we say that a Lévy process with generating triplet (a, σ, U) is of

type A if $\sigma = 0$, $U(\mathbb{R}) < \infty$,

type B if $\sigma = 0$, $U(\mathbb{R}) = \infty$ and $\int_{|x| \leq 1} |x| U(\mathrm{d}x) < \infty$,

type C if $\sigma \neq 0$ or $\int_{|x| \leq 1} |x| U(\mathrm{d}x) = \infty$.

In terms of Cont and Tankov (2003), a process of type A is of *finite activity* (i.e. finitely many jumps), a process of type B is of *infinite activity* (i.e. infinitely many jumps on compacts), both are of finite variation. A process of type C is of *infinite variation* (and infinite activity).

Subordinators Subordinators form a sub-class of Lévy processes. A *subordinator* is defined as a non-decreasing Lévy process – its sample paths are almost surely non-decreasing. This yields that every subordinator has a finite variation, thus it can be of type A or B.

Proposition 2.7. *Let $(L_t : t \geq 0)$ be a Lévy process with generating triplet (a, σ, U). Then the following statements are equivalent:*

(1) L_t is a subordinator.

(2) $a_0 \geq 0$, $\sigma = 0$, $U((-\infty, 0]) = 0$ and

$$\int_0^\infty (x \wedge 1) U(\mathrm{d}x) < \infty.$$

Proof. See (Cont and Tankov, 2003, Proposition 3.10) □

The second condition means that L_t has non-negative drift, no Gaussian part and only positive jumps of finite variation. The characteristic function of a subordinator (L_t) with generating triplet (a, σ, U) can be expressed as

$$\Psi_{L_1}(u) = \exp\left[\mathrm{i}a_0 u + \int_0^\infty \left(e^{\mathrm{i}ux} - 1\right) U(\mathrm{d}x)\right], \qquad u \in \mathbb{R}. \tag{2.9}$$

2.4 Examples of Lévy processes

In the next paragraphs we give some examples of Lévy processes that will be used later.

2.4.1 Poisson process

A Poisson distribution with parameter λ, denoted by $\mathrm{Po}(\lambda)$, is a discrete probability distribution on $\mathbb{N}$. If X has a Poisson distribution, then

$$\mathbb{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}, \ k \in \mathbb{N}.$$

The characteristic function is given by

$$\Psi_X(u) = \exp\left\{\lambda(e^{\mathrm{i}u} - 1)\right\}.$$

A stochastic process $(X_t : t \geq 0)$ on $\mathbb{R}$ is a *Poisson process* with parameter λ, if it is a Lévy process and, for $t \geq 0$, X_t has a Poisson distribution with mean λt.

The Poisson process is a pure jump process with jump size always equal to 1. This means that the generating triplet of a Poisson distribution is given by $(0, 0, \lambda\delta(1))$, where $\delta(1)$ denotes the Dirac measure at point 1, i.e. a measure with a mass 1 concentrated at point 1. Poisson process is of type A, and moreover a subordinator.

2.4.2 Compound Poisson process

We say that a distribution μ on $\mathbb{R}$ is *compound Poisson* if, for some λ and for some distribution σ on $\mathbb{R}$ with $\sigma(\{0\}) = 0$, we have

$$\hat{\mu}(z) = \exp\left\{\lambda(\hat{\sigma}(z) - 1)\right\}, \qquad z \in \mathbb{R}.$$

Let $\lambda > 0$ and let σ be a distribution in $\mathbb{R}$ with $\sigma(\{0\}) = 0$. A stochastic process $(X_t : t \geq 0)$ on $\mathbb{R}$ is a *compound Poisson process* associated with λ and σ if it is a Lévy process and, for $t > 0$, X_t has a compound Poisson distribution, i.e., its characteristic function is given by

$$\Psi_{X_t}(z) = \exp\left\{\lambda t(\hat{\sigma}(z) - 1)\right\} = \exp\left\{t \int_{\mathbb{R}} \left(e^{iux} - 1\right) \nu(dx)\right\}, \tag{2.10}$$

where $\nu(dx) = \lambda\sigma(dx)$. The construction of such process is as follows. Let $(N_t : t \geq 0)$ be a Poisson process with parameter $\lambda > 0$ and let $Y_1, Y_2, \ldots$ be i.i.d. random variables independent of (N_t), having a probability distribution σ with no atom at zero. Then the process

$$X_t = \sum_{i=1}^{N_t} Y_i$$

is a compound Poisson process.

The compound Poisson process is a pure jump process. The law determining the size of the jumps has a distribution σ with no mass in $\{0\}$ and the intensity of jumps is λ. The generating triplet of this Lévy process is given by

$$\left(\lambda \int_{-1}^{1} x\sigma(dx), 0, \lambda\sigma(dx)\right).$$

Compound Poisson process is a process of type A. In the case that the distribution σ is concentrated on $\mathbb{R}^+$, then it is a subordinator.

The significance of compound Poisson processes for Lévy processes arrises with the following result.

Proposition 2.8. *Every infinitely divisible distribution is the limit of a sequence of compound Poisson distributions.*

Proof. Consider an arbitrary infinitely divisible probability measure μ and choose $t_n \downarrow 0$ arbitrarily. Define μ_n by

$$\hat{\mu}_n(z) = \exp\left(t_n^{-1}(\hat{\mu}(z)^{t_n} - 1)\right) = \exp\left(t_n^{-1} \int_{\mathbb{R}\setminus\{0\}} \left(e^{izx} - 1\right) \mu^{t_n}(dx)\right).$$

The distribution μ_n is compound Poisson. Note that

$$\hat{\mu}_n(z) = \exp\left[t_n^{-1}\left(e^{t_n \log \hat{\mu}(z)} - 1\right)\right] = \exp\left[t_n^{-1}(t_n \log \hat{\mu}(z) + O(t_n^2))\right]$$

for each z as $n \to \infty$. Hence $\hat{\mu}_n(z) \to e^{\log \hat{\mu}(z)} = \hat{\mu}(z)$. $\square$

2.4.3 Brownian motion

A stochastic process $(X_t : t \geq 0)$ on $\mathbb{R}$ is a *Brownian motion* or a *Wiener process* if it is a Lévy process and if

(1) for $t > 0$, X_t has a Gaussian distribution with mean 0 and variance t,

(2) there is $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ such that, for every $\omega \in \Omega_0$, $X_t(\omega)$ is continuous in t.

The characteristic function of a Brownian motion $(X_t : t \geq 0)$ is given by

$$\Psi_{X_t}(u) = \exp\left\{-\frac{1}{2}\sigma t u^2\right\}.$$

Brownian motion is a continuous Lévy process with no jump part. Its generating triplet is given by $(0, \sigma, 0)$. It is a process of type C.

2.4.4 Generalized inverse Gaussian process

The generalized inverse Gaussian distribution $\mathrm{GIG}(\lambda,\delta,\gamma)$ is a distribution on $\mathbb{R}^+$ with a density function of the form

$$p_{GIG}(x;\lambda,\delta,\gamma)=\left(\frac{\gamma}{\delta}\right)^{\lambda}\frac{1}{2K_\lambda(\delta\gamma)}x^{\lambda-1}\exp\left\{-\frac{1}{2}(\delta^2x^{-1}+\gamma^2x)\right\},\qquad x\in\mathbb{R}^+ \tag{2.11}$$

where K_λ denotes the modified Bessel function of the third kind with index λ, see Appendix A for more information about the Bessel functions. The parameter space is given by

$$\begin{array}{lll}\delta\ge 0,\ \gamma>0 & \text{if} & \lambda>0\\ \delta>0,\ \gamma>0 & \text{if} & \lambda=0\\ \delta>0,\ \gamma\ge 0 & \text{if} & \lambda<0.\end{array}$$

The mean of the GIG distribution is given by

$$\mathrm{E}[\mathrm{GIG}]=\frac{\delta}{\gamma}\frac{K_{\lambda+1}(\delta\gamma)}{K_\lambda(\delta\gamma)}.$$

Barndorff-Nielsen and Halgreen (1977) and Grosswald (1976) showed that a random variable X having GIG distribution is infinitely divisible and its characteristic function is given by

$$\Psi_{GIG}(u)=\left(1-2\mathrm{i}\frac{u}{\gamma^2}\right)^{\lambda/2}\frac{K_\lambda\left(\delta\gamma\sqrt{1-2\mathrm{i}\frac{u}{\gamma^2}}\right)}{K_\lambda(\delta\gamma)},\qquad \mathrm{Re}(u)<\frac{\gamma^2}{2}. \tag{2.12}$$

A stochastic process $(X_t : t\ge 0)$ on $\mathbb{R}^+$ is a *generalized inverse Gaussian process* if it is a Lévy process where the increment over the time interval $[s,s+t]$, $s,t\ge 0$, has a characteristic function

$$\Big[\Psi_{GIG}(u)\Big]^t.$$

GIG process is a pure jump process with solely positive jumps, hence a subordinator. In the context of Section 2.3. it is a process of type B, of infinite activity with finite variation. Its generating triplet is given by

$$\Big(b_{GIG},0,u_{GIG}(x)\mathrm{d}x\Big),$$

where

$$\begin{aligned} u_{GIG}(x)&=x^{-1}\exp\left(-\frac{1}{2}\gamma^2x\right)\left[\delta^2\int_0^\infty e^{-x\zeta}g_\lambda(2\delta^2\zeta)\mathrm{d}\zeta+\max\{0,\lambda\}\right],\qquad x\in\mathbb{R}^+,\\ b_{GIG}&=\frac{\delta}{\gamma}\frac{K_{\lambda+1}(\delta\gamma)}{K_\lambda(\delta\gamma)}-\int_1^\infty x\,u_{GIG}(x)\mathrm{d}x,\end{aligned} \tag{2.13}$$

and where

$$g_\lambda(x)=\Big[(\pi^2/2)x\left\{J^2_{|\lambda|}(\sqrt{x})+Y^2_{|\lambda|}(\sqrt{x})\right\}\Big]^{-1}.$$

and J_ν and Y_ν are Bessel functions (see also the Appendix A).

We will in particular use the following 2 special cases of GIG processes, the IG (inverse Gaussian) process and the Γ (Gamma) process.

2.4.5 Inverse Gaussian process

The inverse Gaussian distribution $\mathrm{IG}(\delta, \gamma)$ is a special case of GIG distribution with $\lambda = -1/2$. It is concentrated on $\mathbb{R}^+$ and has probability density

$$p_{IG}(x; \delta, \gamma) = \frac{1}{\sqrt{2\pi}} \delta e^{\delta\gamma} x^{-3/2} \exp\left\{-\frac{1}{2}(\delta^2 x^{-1} + \gamma^2 x)\right\}, \qquad x \in \mathbb{R}^+. \tag{2.14}$$

The characteristic function takes the form

$$\Psi_{IG}(u) = \exp\left\{\delta\gamma(1 - \sqrt{1 - 2\mathrm{i}u/\gamma^2})\right\}, \qquad \mathrm{Re}(u) < \frac{\gamma^2}{2}. \tag{2.15}$$

A stochastic process $(X_t : t \geq 0)$ on $\mathbb{R}^+$ is an *inverse Gaussian process* if it is a Lévy process with increments that are $\mathrm{IG}(\delta, \gamma)$ distributed; it follows that X_t has $\mathrm{IG}(t\delta, \gamma)$ distribution. The generating triplet of inverse Gaussian process equals

$$\Big(b_{IG}, 0, u_{IG}(x)\mathrm{d}x\Big),$$

where

$$\begin{aligned} u_{IG}(x) &= \frac{1}{\sqrt{2\pi}} \delta x^{-3/2} \exp(-\gamma^2 x/2), \qquad x \in \mathbb{R}^+, \\ b_{IG} &= \frac{\delta}{\gamma}(2\Phi(\gamma) - 1). \end{aligned} \tag{2.16}$$

2.4.6 Gamma process

The Gamma distribution $\Gamma(\lambda, \gamma)$ is special case of GIG distribution with $\lambda > 0$, $\delta = 0$ and $\gamma > 0$. It is concentrated on $\mathbb{R}^+$ and its probability density is given by

$$p_\Gamma(x; \lambda, \beta) = \beta^\lambda \frac{1}{\Gamma(\lambda)} x^{\lambda-1} \exp\left(-\beta x\right), \qquad x \in \mathbb{R}^+, \beta = \frac{1}{2}\gamma^2. \tag{2.17}$$

The characteristic function of Gamma distribution is given by

$$\Psi_\Gamma(u) = \left(1 - \mathrm{i}u\frac{1}{\beta}\right)^{-\lambda}, \qquad \mathrm{Re}(u) < \beta.$$

A stochastic process $(X_t : t \geq 0)$ on $\mathbb{R}^+$ is a *Gamma process* if it is a Lévy process with increments that are $\Gamma(\lambda, \gamma)$ distributed, i.e. X_t has $\Gamma(t\lambda, \gamma)$ distribution. The generating triplet of Gamma process equals

$$\Big(b_\Gamma, 0, u_\Gamma(x)\mathrm{d}x\Big),$$

where

$$\begin{aligned} u_\Gamma(x) &= \lambda x^{-1} \exp\left(-\beta x\right), \qquad x \in \mathbb{R}^+, \\ b_\Gamma &= \frac{\lambda}{\beta}\left(1 - \exp(-\beta)\right). \end{aligned} \tag{2.18}$$

2.4.7 Normal inverse Gaussian process

The normal inverse gaussian distribution $\mathrm{NIG}(\alpha, \beta, \mu, \delta)$ is special case of the so called generalized hyperbolic distribution with $\lambda = -1/2$. The probability density is given by
(2.19)

$$p_{NIG}(x; \alpha, \beta, \mu, \delta) = \frac{\alpha\delta}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x-\mu)\right) \frac{K_1(\alpha\sqrt{\delta^2 + (x-\mu)^2})}{\sqrt{\delta^2 + (x-\mu)^2}}, \qquad x \in \mathbb{R}.$$

The characteristic function of NIG distribution is given by

$$\Psi_{NIG}(u) = \exp\left(\delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + \mathrm{i}u)^2}\right) + \mu \mathrm{i}u\right), \qquad \mathrm{Re}(|u+\beta|) < \alpha.$$

A stochastic process $(X_t : t \geq 0)$ on $\mathbb{R}$ is a *Normal inverse Gaussian process* if it is a Lévy process with increments that are $\mathrm{NIG}(\alpha, \beta, \mu, \delta)$ distributed, i.e. X_t has $\mathrm{NIG}(\alpha, \beta, t\mu, t\delta)$ distribution. The generating triplet of NIG process equals

$$\left(b_{NIG}, 0, u_{NIG}(x)\mathrm{d}x\right),$$

where

(2.20)
$$u_{NIG}(x) = \frac{\alpha\delta}{\pi} \exp(\beta x) \frac{K_1(\alpha|x|)}{|x|}, \qquad x \in \mathbb{R},$$
$$b_{NIG} = \mu + \frac{2\alpha\delta}{\pi} \int_0^1 \sinh(\beta x) K_1(\alpha x)\mathrm{d}x.$$

The NIG process can be constructed as time changed Brownian motion. Let $(W_t^\beta : t \geq 0)$ be a Brownian motion with drift β and let $(I_t : t \geq 0)$ be an inverse Gaussian process with parameters δ and γ. Then the process defined as

$$X_t = W_{I_t}^\beta + \mu t$$

is NIG process with parameters $\alpha, \beta, \mu, \delta$, where $\alpha = \sqrt{\beta^2 + \gamma^2}$.

Remark 2.9. A review of the properties of GIG and GH distributions may be found in Eberlein (2000) or Schoutens (2003).

Chapter 3

Stochastic integration w.r.t. a Lévy process

In the next Chapters we will study processes that generalizes the diffusion processes by way of replacing the Brownian motion with a Lévy process. Thus we have to define the stochastic integral with respect to a general Lévy process

$$\int_0^t f(s)\mathrm{d}L_s. \tag{3.1}$$

We use as a reference book Jacod and Shiryaev (2003). We start with introducing a stochastic integral with respect to a semimartingale and then show that (3.1) is a special case of such stochastic integral. Then we prove two theorems that are applications of so called Riemann sum representation.

3.1 Construction of the stochastic integral

First we introduce some notation. Consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$.

Definition 3.1. A process M is said to be a local martingale if it is $\mathbb{F}$-adapted and such that the stopped process $M_{t\wedge\tau_n}$ is a martingale for stopping times $\tau_n \uparrow \infty$. We denote by $\mathcal{L}$ the set of all local martingales M such that $M_0 = 0$. We denote by $\mathcal{V}$ the set of all real-valued processes A that are càdlàg (right continuous with left limits), adapted, of finite variation over each finite interval $[0, t]$ and such that $A_0 = 0$.

Definition 3.2. A *semimartingale* is a stochastic process of the form

$$X = X_0 + M + A,$$

where X_0 is finite-valued and $\mathcal{F}_0$ measurable, $M \in \mathcal{L}$ and $A \in \mathcal{V}$.

The first step of the definition or rather construction of a stochastic integral with respect to semimartingales is done for the following class of simple processes.

Definition 3.3. We denote by $\mathcal{E}$ the set of all processes H of the form

$$\begin{cases} \text{either} & H = Y\mathbf{1}_{\{0\}}, \ \ Y \text{ bounded, } \mathcal{F}_0\text{-measurable} \\ \text{or} & H = Y\mathbf{1}_{(r,s]}, \ \ r < s, \ Y \text{ bounded, } \mathcal{F}_r\text{-measurable.} \end{cases}$$

Definition 3.4. Let $H \in \mathcal{E}$ and X be a semimartingale. We define a mapping $H \mapsto H \cdot X$ by

$$H \cdot X_t = \begin{cases} 0 & \text{if } H = Y\mathbf{1}_{\{0\}} \\ Y(X_{s\wedge t} - X_{r\wedge t}) & \text{if } H = Y\mathbf{1}_{(r,s]}. \end{cases} \tag{3.2}$$

We use the following equivalent notation

$$H \cdot X_t = \int_0^t H_s \mathrm{d}X_s = \int_{(0,t]} H_s \mathrm{d}X_s.$$

We define the *ucp* convergence.

Definition 3.5. A sequence of processes $(H^n : n \in \mathbb{N})$ converges to a process H *uniformly on compacts in probability* (abbreviated by $\xrightarrow{ucp}$) if, for each $t > 0$

$$\sup_{0\le s\le t} |H^n_s - H_s| \xrightarrow{P} 0.$$

For the following nontrivial and basic result about the extension of the stochastic integral to locally bounded processes we refer to (Jacod and Shiryaev, 2003, Chapter I, Theorem 4.31).

Theorem 3.6. *Let X be a semimartingale. The mapping $H \mapsto H \cdot X$ defined on $\mathcal{E}$ by* (3.2) *has an extension, still denoted by $H \mapsto H \cdot X$ (and we call $H \cdot X$* stochastic integral of H with respect to X*) to the space of all locally bounded predictable processes H, with the following properties:*

(1) $H \cdot X$ is a càdlàg adapted process,

(2) $H \mapsto H \cdot X$ is linear, up to evanescence,

(3) if a sequence $(H^n : n \in \mathbb{N})$ of predictable processes converges pointwise to a limit H, and if $|H^n| \le K$ where K is a locally bounded predictable process, then $H^n \cdot X_t \to H \cdot X_t$ in measure for all $t \in \mathbb{R}^+$.

Moreover this extension is unique, up to evanescence (i.e. if $H \mapsto \alpha(H)$ is another extension with the same properties, then $\alpha(H)$ and $H \cdot X$ are indistinguishable), and in (3) above $H^n \cdot X \xrightarrow{ucp} H \cdot X$.

By specifying the integrator as a locally square integrable martingale (we denote the class of all locally square integrable martingales by $\mathcal{H}^2_{\mathrm{loc}}$), we are able to enlarge the class of integrands from the class of locally bounded predictable processes to a class of all predictable processes that are locally square integrable with respect to the norm induced by the quadratic variation of the locally square integrable martingale integrator (we denote such a class by $L^2_{\mathrm{loc}}(X)$, for $X \in \mathcal{H}^2_{\mathrm{loc}}$). We refer to (Jacod and Shiryaev, 2003, Chapter I, Theorem 4.40) for the following precise form of this extension.

Theorem 3.7. *Let $X \in \mathcal{H}^2_{\mathrm{loc}}$. The mapping $H \mapsto H \cdot X$ (defined either on $\mathcal{E}$ by* (3.2) *of for all locally bounded predictable process H by the Theorem 3.6) has a further extension to the set $L^2_{\mathrm{loc}}(X)$, still denoted by $H \mapsto H \cdot X$, which meets the conditions (1), (2) from the Theorem 3.6 and*

(3) if a sequence (H^n) of predictable processes converges pointwise to a limit H and $|H^n| \leq K$ for some $K \in L^2_{\text{loc}}(X)$, then $H^n \cdot X \xrightarrow{ucp} H \cdot X$.

Moreover this extension is unique (up to evanescence), and we have:

(a) $H \cdot X \in \mathcal{H}^2_{\text{loc}}$.

(b) $H \cdot X \in \mathcal{H}^2$ if and only if $H \in L^2(X)$.

(c) If $X, Y \in \mathcal{H}^2_{\text{loc}}$ and $H \in L^2_{\text{loc}}(X)$ and $K \in L^2_{\text{loc}}(Y)$, then

$$\langle H \cdot X, K \cdot Y \rangle = (HK) \cdot \langle X, Y \rangle. \tag{3.3}$$

To make these results constructive, we show that the stochastic integral of a predictable càg process may be approximated by Riemann sums. First we introduce a notation. Let Δ be a subdivision of the nonnegative real line $\mathbb{R}^+$, $0 = t_0 < t_1 < \cdots < \infty$. The Δ-*Riemann approximant* of $H \cdot X$ is the process $\Delta(H \cdot X)$ defined by

$$\Delta(H \cdot X)_t = \sum_{t_j \in \Delta} H_{t_j}(X_{t_{j+1} \wedge t} - X_{t_j \wedge t}).$$

We will be referring to the following result, which is an adaption of (Jacod and Shiryaev, 2003, Chapter I, Proposition 4.44), as the *Riemann sum representation theorem.*

Theorem 3.8. *Let X be a semimartingale, H be a càg adapted process, and $(\Delta^n : n \in \mathbb{N})$ a sequence of subdivisions of the nonnegative real line $\mathbb{R}^+$, such that $|\Delta^n| = \sup\{|t^n_j - t^n_{j-1}|, t^n_j \in \Delta^n\} \downarrow 0$ for $n \to \infty$. Then*

$$\Delta^n(H \cdot X) \xrightarrow{ucp} H \cdot X. \tag{3.4}$$

Proof. Define a new process H^n by

$$H^n = \sum_{t_j \in \Delta^n} H_{t^n_j} \mathbf{1}_{(t^n_j, t^n_{j+1}]}.$$

Then H^n is predictable, converges pointwise to H, because H is right continuous. Suppose that $K_t = \sup_{s \leq t} |H_s|$. Then K is adapted, càg, locally bounded, and $|H^n| \leq K$. Hence, from the property (3) of the Theorem 3.6

$$H^n \cdot X \xrightarrow{ucp} H \cdot X.$$

Moreover we have

$$H^n \cdot X_t = \sum_{t_j \in \Delta^n} H_{t^n_j}(X_{t^n_{j+1} \wedge t} - X_{t^n_j \wedge t}) = \Delta^n(H \cdot X)_t.$$

This completes the proof. $\square$

3.2 Lévy process as a semimartingale

In this section we show that every Lévy process is a semimartingale. Every semimartingale has a form of sum of three processes: if X is a semimartingale, then

$$X = X_0 + M + A,$$

where X_0 is finite-valued, $\mathcal{F}_0$ measurable process, M is a local martingale and A is an adapted càdlàg process with finite variation. On the other hand we know that every Lévy process has a Lévy-Itô decomposition (2.8): if L is a Lévy process, then

$$L = L^{(0)} + L^{(1)} + L^{(2)} + L^{(3)}$$

where $L_t^{(0)} = at$ is a deterministic drift and hence a process of finite variation.

$$L_t^{(1)} = \sqrt{\sigma} W_t$$

is a Brownian motion and hence a local martingale.

$$L_t^{(2)} = \sum_{s \leq t} |\Delta_s| \mathbf{1}_{\{|\Delta_s| \geq 1\}}$$

is the sum of jumps of the process L that are bigger or equal to one, more precisely a compound Poisson process with jumps bigger or equal to one and hence a process of finite variation, see (Sato, 1999, Theorem 21.9, (i)). Finally

$$L_t^{(3)} = \lim_{\varepsilon \to 0} \Big(\sum_{s \leq t} |\Delta_s| \mathbf{1}_{\{\varepsilon < |\Delta_s| < 1\}} - t \int_{\varepsilon < |x| < 1} x U(\mathrm{d}x) \Big)$$

is the compensated sum of jumps of the process L that are smaller than one, more precisely a compensated compound Poisson process and hence a local martingale. If we make the following rearrangement

$$\begin{aligned} A &= L^{(0)} + L^{(2)} \\ M &= L^{(1)} + L^{(3)}, \end{aligned}$$

then L becomes a semimartingale indeed.

3.3 Consequences of Riemann sum representation

In this Section we prove a fundamental result that will be referred to as *Key theorem.* This result is the conditional form of (Eberlein and Raible, 1999, Lemma 3.1). Next we show that the integrated Lévy process has independent increments. Both these Theorems are a consequence of the Riemann sum representation.

Theorem 3.9 (Key theorem)**.** *Let L_t be a Lévy process such that* $\mathrm{Dom}(\psi_{L_1})$ *is open and contains a translate of the imaginary axis. Let $f : \mathbb{R}^+ \to \mathbb{C}$ be a measurable, left-continuous function with limits from the right, such that* $\mathrm{Im}(f) \subseteq \mathrm{Dom}(\psi_{L_1})$*. Then, for any $0 \leq s < t$, we have*

$$\mathrm{E}\left[\exp\left(\int_s^t f(u)\mathrm{d}L_u\right) |\mathcal{F}_s\right] = \exp\left(\int_s^t \theta_{L_1}(f(u))\mathrm{d}u\right). \tag{3.5}$$

Proof. Step 1: Consider an arbitrary partititon $s = t_0 < t_1 < \cdots < t_N = t$ of the time interval $[s,t]$ with mesh $\Delta = \sup_{j=1,\cdots,N} |t_j - t_{j-1}|$ and denote $\tau_j = t_j - s$, for $j = 0, \cdots, N$. We get

$$\begin{aligned} \mathrm{E}\left[\exp\left(\sum_{j=0}^{N-1} f(t_i)(L_{t_{j+1}} - L_{t_j})\right)|\mathcal{F}_s\right] &= \prod_{j=0}^{N-1} \mathrm{E}\left[\exp\left(f(t_j)(L_{t_{j+1}} - L_{t_j})\right)|\mathcal{F}_s\right] = \\ &= \prod_{j=0}^{N-1} \mathrm{E}\left[\exp\left(f(\tau_j + s)(L_{\tau_{j+1}+s} - L_{\tau_j+s})\right)|\mathcal{F}_s\right]. \end{aligned}$$

Here we use the restarting argument. We define a new Lévy process $\hat{L}_\tau = L_{\tau+s} - L_s$ that is independent of $\mathcal{F}_s$. Then we can write

$$\prod_{j=0}^{N-1} \mathrm{E}\left[\exp\left(f(\tau_j + s)(L_{\tau_{j+1}+s} - L_{\tau_j+s})\right)|\mathcal{F}_s\right] = \prod_{j=0}^{N-1} \mathrm{E}\left[\exp\left(f(\tau_j + s)(\hat{L}_{\tau_{j+1}} - \hat{L}_{\tau_j})\right)\right] =$$

and we use the stationary increments property to show

$$= \prod_{j=0}^{N-1} \mathrm{E}\left[\exp\left(f(\tau_j + s)\hat{L}_{\tau_{j+1}-\tau_j}\right)\right].$$

Using the property of the moment generating function of a Lévy process

$$\psi_{L_t}(u) = [\psi_{L_1}(u)]^t$$

we obtain

$$\begin{aligned} \prod_{j=0}^{N-1} \mathrm{E}\left[\exp\left(f(\tau_j + s)L_{\tau_{j+1}-\tau_j}\right)\right] &= \prod_{j=0}^{N-1} \left[\psi_{L_1}\left(f(\tau_j + s)\right)\right]^{(\tau_{j+1}-\tau_j)} = \\ &= \prod_{j=0}^{N-1} \exp\left(\theta_{L_1}\left(f(\tau_j + s)\right)\ (\tau_{j+1} - \tau_j)\right) = \\ &= \exp\left(\sum_{j=0}^{N-1} \theta_{L_1}(f(\tau_j + s))(\tau_{j+1} - \tau_j)\right). \end{aligned}$$

Step 2: With mesh $\Delta \downarrow 0$ we have, using the Riemann sum representation theorem 3.8

$$\mathrm{E}\left[\exp\left(\int_s^t f(u)\mathrm{d}L_u\right)|\mathcal{F}_s\right] = \mathrm{E}\left[\exp\left(\mathbb{P} - \lim_{|\Delta|\downarrow 0} \sum_{j=0}^{N-1} f(t_j)(L_{t_{j+1}} - L_{t_j})\right)|\mathcal{F}_s\right]$$

We apply the dominated convergence theorem (using the fact that f is bounded and exp is continuous), and use the result of Step 1

$$\begin{aligned} &= \lim_{|\Delta|\downarrow 0} \mathrm{E}\left[\exp\left(\sum_{j=0}^{N-1} f(t_j)(L_{t_{j+1}} - L_{t_j})\right)|\mathcal{F}_s\right] \\ &= \lim_{|\Delta|\downarrow 0} \exp\left(\sum_{j=0}^{N-1} \theta_{L_1}(f(\tau_j + s))(\tau_{j+1} - \tau_j)\right) \end{aligned}$$

Again by the continuity of exp, we can rewrite this into

$$= \exp\Big(\lim_{|\Delta|\downarrow 0} \sum_{j=0}^{N-1} \theta_{L_1}(f(\tau_j + s))(\tau_{j+1} - \tau_j) \Big)$$

Using the fact that $\theta_{L_1} \circ f$ is Riemann-integrable on $[0, t-s]$ and by substitution $u = \tau + s$ we obtain

$$= \exp\left(\int_0^{t-s} f(\tau + s)\mathrm{d}\tau \right) = \exp\left(\int_s^t f(u)\mathrm{d}u \right).$$

□

Theorem 3.10 (Independent increments of integrated Lévy process)**.** *Let φ be a càg real-valued function and $(L_t : t \geq 0)$ a Lévy process. Define*

$$X_t = \int_0^t \varphi(s)\mathrm{d}L_s.$$

Then $(X_t : t \geq 0)$ is a process with independent increments such that $X_t - X_s$ is independent of $\mathcal{F}_s$ for any $0 \leq s \leq t$.

Proof. Let $s \leq t$ and $(\Delta^n : n \in \mathbb{N})$ a sequence of subdivisions of the time interval $[s, t]$ with $|\Delta^n| = \sup\{|t^n_{j+1} - t^n_j| : t^n_j \in \Delta^n\} \downarrow 0$ for $n \to \infty$.
Step 1: First we show that $X_t - X_s$ is independent of $\mathcal{F}_s$. Using the Riemann sum representation theorem

$$\begin{aligned} X_t - X_s = \int_s^t \varphi(u)\mathrm{d}L_u &= \mathbb{P} - \lim_{n\to\infty} \sum_{t^n_j \in \Delta^n} \varphi(t^n_j)(L_{t^n_{j+1}} - L_{t^n_j}) = \\ &= \mathbb{P} - \lim_{n\to\infty} \sum_{t^n_j \in \Delta^n} \varphi(t^n_j)((L_{t^n_{j+1}} - L_s) - (L_{t^n_j} - L_s)). \end{aligned}$$

We define a new Lévy process by $\hat{L}_{t-s} = L_t - L_s$ and we refer to the restarting argument in the proof of the Theorem 3.9 to obtain

$$X_t - X_s = \mathbb{P} - \lim_{n\to\infty} \sum_{t^n_j \in \Delta^n} \varphi(\tau^n_j + s)(\hat{L}_{t^n_{j+1}-s} - \hat{L}_{t^n_j - s}).$$

The summands of the last term are random variables that are indendent of $\mathcal{F}_s$. Hence also $X_t - X_s$ is independent of $\mathcal{F}_s$.
Step 2: As a second step we show that $X_t - X_s$ and X_s are independent. For that we need to show that

$$\mathrm{E}[(X_t - X_s)X_s] = 0.$$

Consider a sequence of subdivisions $(\bar{\Delta}^n : n \in \mathbb{N})$ of the time interval $[0, s]$ with $|\bar{\Delta}^n| = \sup\{|t^n_{k+1} - t^n_k| : t^n_k \in \bar{\Delta}^n\} \downarrow 0$ for $n \to \infty$. Using the Riemann sum representation theorem

$$\begin{aligned} \mathrm{E}[(X_t - X_s)X_s] &= \mathrm{E}\left[\int_s^t \varphi(u)\mathrm{d}L_s \int_0^s \varphi(v)\mathrm{d}L_v \right] = \\ &= \mathrm{E}\Big[\mathbb{P} - \lim_{n\to\infty} \sum_{t^n_j \in \Delta^n} \varphi(t^n_j)(L_{t^n_{j+1}} - L_{t^n_j}) \cdot \mathbb{P} - \lim_{n\to\infty} \sum_{t^n_k \in \bar{\Delta}^n} \varphi(t^n_k)(L_{t^n_{k+1}} - L_{t^n_k}) \Big] = \end{aligned}$$

Here we apply the dominated convergence theorem to obtain

$$\begin{aligned}
\mathrm{E}[(X_t - X_s)X_s] &= \lim_{n\to\infty} \mathrm{E}\Bigg[\sum_{t_j^n\in\Delta^n}\sum_{t_k^n\in\bar{\Delta}^n} \varphi(t_j^n)\varphi(t_k^n)(L_{t_{j+1}^n} - L_{t_j^n})(L_{t_{k+1}^n} - L_{t_k^n})\Bigg] = \\
&= \lim_{n\to\infty} \sum_{t_j^n\in\Delta^n}\sum_{t_k^n\in\bar{\Delta}^n} \varphi(t_j^n)\varphi(t_k^n)\mathrm{E}\left[(L_{t_{j+1}^n} - L_{t_j^n})(L_{t_{k+1}^n} - L_{t_k^n})\right]
\end{aligned}$$

From the condition (1) of the defnition of Lévy process and using the fact that $t_k < t_{k+1} \leq s \leq t_j < t_{j+1}$ for all $t_k \in \bar{\Delta}^n$ and $t_j \in \Delta^n$ it follows that

$$\mathrm{E}\left[(L_{t_{j+1}^n} - L_{t_j^n})(L_{t_{k+1}^n} - L_{t_k^n})\right] = 0$$

for every $t_k \in \bar{\Delta}^n$ and $t_j \in \Delta^n$. Hence we conclude

$$\mathrm{E}[(X_t - X_s)X_s] = 0.$$

□

Chapter 4

Ornstein–Uhlenbeck type processes

In this chapter we review pertinent facts about selfdecomposable processes, their properties and relation with Lévy processes. As typical examples of selfdecomposable distributions we discuss processes of Ornstein–Uhlenbeck type. The discussion follows Sato (1999) and Barndorff-Nielsen and Shephard (2000).

4.1 Selfdecomposable distributions

Definition 4.1. We call a probability measure μ on $\mathbb{R}$ *selfdecomposable*, if, for any $\lambda > 0$, there is a probablity measure ϱ_λ such that

$$\hat{\mu}(z) = \hat{\mu}(e^{-\lambda} z)\hat{\varrho}_\lambda(z), \qquad \text{for all } z \in \mathbb{R}. \tag{4.1}$$

Selfdecomposable distributions furnish a subclass of the infinititely divisible distributions as follows.

Proposition 4.2. *If a probability distribution μ is selfdecomposable, then it is also infinitely divisible, and, for any $\lambda > 0$, the probability measure ϱ_λ is uniquely determined and infinitely divisible.*

Proof. See (Sato, 1999, Proposition 15.5.). □

Criteria for selfdecomposability A further important characterization of the class of selfdecomposable distributions as a subclass of the set of all infinitely divisible distributions in terms of the Lévy measure is the following equivalence. We shall use the following notation. Let U be a Lévy measure. We denote the tail masses of the measure U by

$$U^-(x) = U((-\infty, x]) \text{ and } U^+(x) = U([x, \infty)).$$

Proposition 4.3. *Let $U(\mathrm{d}x)$ denote the Lévy measure of an infinitely divisible probability measure μ on $\mathbb{R}$. Then the following statements are equivalent:*

(1) μ is selfdecomposable,

(2) The functions on $\mathbb{R}^+$ given by $U^+(e^s)$ and $U^-(e^s)$ are both convex,

(3) U is of the form $U(\mathrm{d}x) = u(x)\mathrm{d}x$ with $\bar{u}(x) = |x|u(x)$ increasing on $(-\infty, 0)$ and decreasing on $(0, \infty)$.

Proof. See (Barndorff-Nielsen and Shephard, 2000, Theorem 4.1). □

Example 4.4. GIG distributions are selfdecomposable.

Proof. We use the implication (3) $\Rightarrow$ (1) in the Proposition 4.3 and show that $\bar{u}(x) = |x|u(x)$, with $u(x)$ given by (2.13), is increasing on $(-\infty, 0)$ and decreasing on $(0, \infty)$. We have

$$\bar{u}(x) = \begin{cases} xu(x) = \delta^2 e^{-\frac{\gamma^2}{2}x} \int_0^\infty e^{-x\zeta} g_\lambda(2\delta^2\zeta)\mathrm{d}\zeta + e^{-\frac{\gamma^2}{2}x} \max\{0, \lambda\}, & x > 0, \\ -xu(x) = 0, & x \leq 0, \end{cases} \tag{4.2}$$

with $\lambda \in \mathbb{R}$ and $\delta, \gamma \in \mathbb{R}_0^+$ not simultaneously equal to 0 and

$$g_\lambda(z) = \left[(\pi^2/2)z\left\{J^2_{|\lambda|}(\sqrt{z}) + Y^2_{|\lambda|}(\sqrt{z})\right\}\right]^{-1}, \qquad z > 0.$$

The function $\bar{u}(x)$ is increasing on $(-\infty, 0)$. The exponential $\exp(-\gamma^2 x/2)$ is decreasing on $(0, \infty)$, hence the problem reduces to showing that the integral in (4.2) is decreasing. This holds if and only if the derivative of the integral is negative. The derivative under the integral sign gives a negative sign and since $g_\lambda(z) > 0$ for all $z \in (0, \infty)$ we conclude that $\bar{u}(x)$ is decreasing on $(0, \infty)$. □

Representation of selfdecomposable distributions A random variable X is selfdecomposable if and only if it has a representation given by the following result.

Proposition 4.5. *A random variable X is selfdecomposable if and only if*

$$X = \int_0^\infty e^{-t}\mathrm{d}L_t,$$

where $(L_t : t \geq 0)$ is a Lévy process.

Proof. See (Jurek and Vervaat, 1983, Theorem 3.2.) □

4.2 Ornstein–Uhlenbeck type processes

This subsection reviews the construction of Ornstein-Uhlenbeck type processes following the paper of Barndorff-Nielsen et al. (1998).

Definition 4.6. Given a Lévy process $(L_t : t \geq 0)$ generated by (b, τ, U), real constants $a > 0$, $\sigma > 0$ and a random variable X_0 independent of (L_t), we define a new stochastic process $(X_t : t \geq 0)$ as

$$X_t = e^{-at}X_0 + \int_0^t \sigma e^{-a(t-s)}\mathrm{d}L_s. \tag{4.3}$$

We call $(X_t : t \geq 0)$ an Ornstein-Uhlenbeck type process generated by (b, τ, U, a, σ). This process verifies the linear "stochastic integral equation"

$$X_t = X_0 - a\int_0^t X_s\mathrm{d}s + \sigma L_t. \tag{4.4}$$

Its local behaviour is described via a stochastic differential equation as the sum of a linear damping term and a random term described by the increments of L_t:

$$\mathrm{d}X_t = -aX_t\mathrm{d}t + \sigma\mathrm{d}L_t. \tag{4.5}$$

We define an *Ornstein-Uhlenbeck type process with drift* (generated by $(b, \tau, U, a, \sigma, \vartheta)$) as a process $(X_t : t \geq 0)$ that satisfes the following SDE

$$\mathrm{d}X_t = (\vartheta(t) - aX_t)\mathrm{d}t + \sigma\,\mathrm{d}L_t.$$

The Lévy process (L_t) is called the *background driving Lévy process* (BDLP) corresponding to the process (X_t).

Remark 4.7. It is easy to see that an OU type process is a Markov process. Sato (1999) defines OU type process as a Markov process by specifying the transition function.

Construction of Ornstein-Uhlenbeck type process We can rewrite the OU type process (4.3) as

$$X_{t+u} = e^{-au}X_t + \int_0^u \sigma e^{-a(u-s)}\mathrm{d}L_{t+s} \overset{d}{=} e^{-au}X_t + \int_0^u \sigma e^{-a(u-s)}\mathrm{d}L_s.$$

Denote the distribution of X_t by μ and the distribution of the stochastic integral

$$\int_0^t \sigma e^{-a(t-s)}\mathrm{d}L_s$$

by ϱ_t. It follows that X is a stationary process if and only if

$$\widehat{\mu}(z) = \widehat{\mu}(e^{-au}z)\widehat{\varrho}_u(z) \tag{4.6}$$

holds for all $u > 0$, i.e. if and only if X_t is selfdecomposable. The characteristic function $\widehat{\varrho}_t$ may be rewritten in the following form, by application of the Key theorem 3.9.

$$\begin{aligned}\widehat{\varrho}_t(z) &= \mathrm{E}\Big[\exp\Big(\int_0^t \mathrm{i}z\sigma e^{-a(t-s)}\mathrm{d}L_s\Big)\Big] = \\ &= \exp\Big(\int_0^t \Theta_{L_1}(z\sigma e^{-a(t-s)})\mathrm{d}s\Big) = \\ &= \exp\Big(\int_0^t \Theta_{L_1}(z\sigma e^{-as})\mathrm{d}s\Big).\end{aligned}$$

By substitution $w = z\sigma e^{-as}$ we get

$$\widehat{\varrho}_t(z) = \exp\Big(\int_{z\sigma e^{-at}}^{z\sigma} \Theta_{L_1}(w)a^{-1}w^{-1}\mathrm{d}w\Big). \tag{4.7}$$

It follows from (4.6) and (4.7) that

$$\widehat{\mu}(z) = \lim_{t\to\infty}\widehat{\varrho}_t(z) = \exp\Big(a^{-1}\int_0^{z\sigma} \Theta_{L_1}(w)w^{-1}\mathrm{d}w\Big). \tag{4.8}$$

The convergence of the integral on the right hand side of (4.8) is thus a necessary condition for X to be stationary. If we want X_t to have a distribution μ we have to choose a Lévy process satysfying

$$\int_0^z |\Theta_{L_1}(w)|w^{-1}\mathrm{d}w < \infty, \qquad \text{for all } z > 0.$$

We can choose L by choosing an appropriate characteristic function in the following way.

Lemma 4.8. *Suppose that μ is selfdecomposable distribution with characteristic function that is differentiable for $z \neq 0$ and suppose that $z\kappa'(z)$ can be defined at zero by continuity, where $\kappa(z) = \log \hat{\mu}(z)$. Then $\exp(z\kappa'(z))$ is an infinitely divisible characteristic function.*

Sketch of proof. See Barndorff-Nielsen et al. (1998). From the fact that μ is selfdecomposable we can show that

$$[\hat{\mu}(z)/\hat{\mu}(sz)]^{\lambda(1-s)^{-1}}$$

is a characteristic function for all $s \in [0, 1)$ and $\lambda > 0$ and from continuity also its limit for $s \to 1$ is a characteristic function. The limit of the logarithm is

$$\lim_{s\to 1} \log[\hat{\mu}(z)/\hat{\mu}(sz)]^{\lambda(1-s)^{-1}} = \lim_{s\to 1} \lambda z \frac{\kappa(z) - \kappa(sz)}{z(1-s)} = \lambda z\kappa'(z),$$

for all $\lambda > 0$ and hence $[\exp(z\kappa'(z))]^{\lambda}$ is a characteristic function for all $\lambda > 0$, and thus $\exp(z\kappa'(z))$ is an infinitely divisible characteristic function. □

For every selfdecomposable distribution μ we may find a Lévy process such that the process X defined by (4.3) is stationary and has distribution given by μ. We summarize the previous development in the following result.

Proposition 4.9. *Let X be selfdecomposable with characteristic function satisfying the conditions of the Lemma 4.8. Then there exists a stationary stochastic process $(X_t : t \geq 0)$ and a Lévy process $(L_t : t \geq 0)$, independent of X_0, such that $X_t \overset{d}{=} X$ and*

$$X_t = e^{-at}X_0 + \int_0^t \sigma e^{-a(t-s)}\mathrm{d}L_s \tag{4.9}$$

for all $a > 0$ and $\sigma > 0$.

***D*–OU type processes** We have shown that for every selfdecomposable distribution there exists a stationary OU type process having this distribution. We will use the following definition introduced in Barndorff-Nielsen and Shephard (2000).

Definition 4.10. If X is an OU type process with marginal distribution D then we say that $(X_t : t \geq 0)$ is a *D–OU type process.* Further, if the BDLP at time 1, i.e. L_1, has a distribution $\tilde{D}$ then we say that $(X_t : t \geq 0)$ is an *OU–$\tilde{D}$ type process.*

Relationships between OU type process and BDLP There is a relationship between the moment generating function of the OU type process X_t and the BDLP L_1 given in the following result.

Proposition 4.11. *Let X be an OU type process and let L be the corresponding BDLP. Then the logarithm of moment generating functions of X_t and L_1 are related by*

$$\theta_{X_t}(z) = \int_0^{z\sigma} \theta_{L_1}(u) a^{-1} u^{-1} \mathrm{d}u \tag{4.10}$$

and

$$\theta_{L_1}(z\sigma) = az \frac{\partial \theta_{X_t}(z)}{\partial z}. \tag{4.11}$$

Proof. X_t is selfdecomposable and it follows from (4.8) that the moment generating function of X_t is

$$\begin{aligned} \psi_{X_t}(\zeta) &= \lim_{t\to\infty} \psi_{I_t}(z) \\ &= \lim_{t\to\infty} \exp\Big(\int_0^t \theta_{L_1}(z\sigma e^{-as})\mathrm{d}s\Big) \\ &= \exp\Big(\int_0^\infty \theta_{L_1}(z\sigma e^{-as})\mathrm{d}s\Big). \end{aligned}$$

where I_t is the stochastic integral

$$I_t = \int_0^t \sigma e^{-a(t-s)} \mathrm{d}L_s.$$

After the substitution $u = z\sigma e^{-as}$ we obtain the first part of the Proposition. The second part of the Proposition follows from the first part by differentiation with respect to z,

$$\frac{\partial \theta_{X_t}(z)}{\partial z} = \theta_{L_1}(z\sigma)\sigma a^{-1} z^{-1} \sigma^{-1}.$$

□

A relationship between the generating triplet of an D–OU type process X and the corresponding BDLP L can be shown.

Proposition 4.12. *Let D be a selfdecomposable distribution with generating triplet (c, υ, V) and $(X_t : t \geq 0)$ be a D–OU type process. Then the generating triplet (b, τ, U) of the BDLP $(L_t : t \geq 0)$ and (c, υ, V) are related in the following way*

$$c = b\frac{\sigma}{a} + \frac{1}{a}\int_{|x|>\sigma^{-1}} \frac{x}{|x|} U(\mathrm{d}x) + \frac{\sigma}{a}\int_{1<|x|\leq\sigma^{-1}} xU(\mathrm{d}x), \tag{4.12}$$

$$\upsilon = \tau\frac{\sigma^2}{2a}, \tag{4.13}$$

$$V(\mathrm{d}x) = \int_0^\infty U(\sigma^{-1}e^{as}\mathrm{d}x)\mathrm{d}s. \tag{4.14}$$

Proof. On one hand we have

$$\Psi_{X_t}(z) = \Psi_{IG}(z) = \exp\left[\mathrm{i}cz - \frac{1}{2}\upsilon^2 z + \int_{\mathbb{R}} \left(e^{\mathrm{i}zx} - 1 - \mathrm{i}zx1_D(x)\right) V(\mathrm{d}x)\right]. \tag{4.15}$$

On the other hand it follows from (4.8) that $\Psi_{X_t}(z) = \lim_{t\to\infty} \Psi_{I_t}(z)$, where

$$I_t = \int_0^t \sigma e^{-a(t-s)} \mathrm{d}L_s.$$

By application of the Key theorem 3.9, we obtain

$$\begin{aligned}
\Psi_{X_t}(z) &= \lim_{t\to\infty} \mathrm{E}[\exp(\mathrm{i}zI_t)] = \\
&= \lim_{t\to\infty} \exp\left(\int_0^t \theta_{L_1}\left(\mathrm{i}z\sigma e^{-a(t-s)}\right) \mathrm{d}s\right) = \\
&= \exp\left(\int_0^\infty \theta_{L_1}\left(\mathrm{i}z\sigma e^{-s}\right) \mathrm{d}s\right) \qquad (4.16)
\end{aligned}$$

From the Lévy-Khintchine formula for the BDLP we know that

$$\theta_{L_1}(z) = bz + \frac{1}{2}\tau z^2 + \int_{\mathbb{R}} \left(e^{zx} - 1 - zx1_D(x)\right) U(\mathrm{d}x). \qquad (4.17)$$

Hence, plugging (4.17) into (4.16), we obtain

$$\begin{aligned}
\Psi_{X_t}(z) &= \exp\left[\int_0^\infty \mathrm{i}bz\sigma e^{-as}\mathrm{d}s - \int_0^\infty \frac{1}{2}\tau z^2\sigma^2 e^{-2as}\mathrm{d}s + \right.\\
&\qquad \left. + \int_0^\infty \int_{\mathbb{R}} \left(\exp(\mathrm{i}z\sigma e^{-as}x) - 1 - \mathrm{i}z\sigma e^{-as}x1_D(x)\right) U(\mathrm{d}x)\mathrm{d}s\right] = \\
&= \exp\left[\mathrm{i}zb\frac{\sigma}{a} - \frac{1}{2}z^2\tau\frac{\sigma^2}{2a} + \int_0^\infty \left(e^{\mathrm{i}zy} - 1 - \mathrm{i}zy1_D(y)\right) \int_{\mathbb{R}} U(\sigma^{-1}e^{as}\,\mathrm{d}x)\mathrm{d}s + \right.\\
&\qquad \left. + \mathrm{i}z \int_0^\infty \int_{\mathbb{R}} \sigma e^{-as}x\left(1_D(\sigma e^{-as}x) - 1_D(x)\right) U(\mathrm{d}x)\mathrm{d}s\right], \qquad (4.18)
\end{aligned}$$

where in the second equation we used the substitution $y = \sigma e^{-as}x$. It stays to compute the last double integral. We have to consider two cases. When $\sigma \in (0, 1]$, then the double integral in (4.18) can be written as

$$\begin{aligned}
\int_0^\infty \int_{1<|x|\le\sigma^{-1}e^{as}} \sigma e^{-as}x\,U(\mathrm{d}x)\mathrm{d}s &= \int_{1<|x|\le\sigma^{-1}} \sigma x\,U(\mathrm{d}x) \int_0^\infty e^{-as}\mathrm{d}s + \\
&\quad + \int_{|x|>\sigma^{-1}} \sigma x \int_{\frac{1}{a}\ln(\sigma|x|)}^\infty e^{-as}\mathrm{d}s\,U(\mathrm{d}x) = \\
&= \int_{1<|x|\le\sigma^{-1}} x\frac{\sigma}{a}U(\mathrm{d}x) + \int_{|x|>\sigma^{-1}} \frac{1}{a}\frac{x}{|x|}U(\mathrm{d}x).
\end{aligned}$$

When $\sigma > 1$ we can write the double integral in (4.18) as

$$\begin{aligned}\int_0^\infty \int_{1<|x|\le\sigma^{-1}e^{as}} \sigma e^{-as} x\, U(\mathrm{d}x)\mathrm{d}s &= -\int_0^{\frac{1}{a}\ln\sigma} \int_{\sigma^{-1}e^{as}<|x|\le 1} \sigma e^{-as} x\, U(\mathrm{d}x)\mathrm{d}s + \\ &\quad + \int_{\frac{1}{a}\ln\sigma}^{\infty} \int_{1<|x|<\sigma^{-1}e^{as}} \sigma e^{-as} x\, U(\mathrm{d}x)\mathrm{d}s = \\ &= -\int_{\sigma^{-1}<|x|\le 1} \sigma x \int_0^{\frac{1}{a}\ln(\sigma|x|)} e^{-as}\mathrm{d}s\, U(\mathrm{d}x) + \\ &\quad + \int_{|x|>1} \sigma x \int_{\frac{1}{a}\ln(\sigma|x|)}^{\infty} e^{-as}\mathrm{d}s\, U(\mathrm{d}x) = \\ &= \int_{|x|>\sigma^{-1}} \frac{1}{a}\frac{x}{|x|} U(\mathrm{d}x) - \int_{\sigma^{-1}<|x|\le 1} \frac{\sigma}{a} x\, U(\mathrm{d}x).\end{aligned}$$

Summarizing the terms corresponding to c, τ and W respectively, we obtain the equalities (4.12) – (4.14). □

Corollary 4.13. *If the Lévy density v of the OU type process X is differentiable, then the Lévy measure U of L_1 has a density u, and v and u are related by*

$$u(\sigma^{-1}x) = a\sigma(-v(x) - xv'(x)). \tag{4.19}$$

Proof. The proof consists of differentiation of the tail measures. Let $x \ge 0$. From (4.14) follows that the tail measure $V^+(x)$ can be written as

$$V([x,\infty)) = \int_0^\infty U(\sigma^{-1}e^{as}[x,\infty))\mathrm{d}s = \int_x^\infty U^+(\sigma^{-1}y)a^{-1}y^{-1}\mathrm{d}y$$

after substitution $y = e^{as}x$. In the same time we have

$$V([x,\infty)) = \int_x^\infty v(y)\mathrm{d}y.$$

Thus

$$v(x) = a^{-1}x^{-1}U^+(\sigma^{-1}x), \qquad \text{for } x \ge 0.$$

In similar way we obtain for $x < 0$

$$v(x) = a^{-1}|x|^{-1}U^-(\sigma^{-1}x), \qquad \text{for } x < 0.$$

By differentiation of v (note that $\left(U^+(\sigma^{-1}x)\right)' = -\sigma^{-1}u(\sigma^{-1}x)$) we obtain that

$$v(x) = -x^{-1}v'(x) - a^{-1}\sigma^{-1}x^{-1}u(\sigma^{-1}x)$$

and by simple rearrangement we obtain the desired result. □

Infinite divisibility of OU type process Every OU type process is infinitely divisible and its Lévy-Khintchine representation is given by the following result.

Proposition 4.14. *Let $a > 0, \sigma \in (0,1)$ and $(L_t : t \geq 0)$ be a Lévy process on $\mathbb{R}$ with generating triplet (b, τ, U). The OU type process generated by (b, τ, U, a, σ) defined by (4.3) is infinitely divisible for every t and has generating triplet (c_t, υ_t, V_t) with*

$$\begin{aligned} c_t &= e^{-at}X_0 + b\frac{\sigma}{a}\left(1 - e^{-at}\right) + \int_0^t \sigma e^{-as}\int_{\mathbb{R}} \mathbf{1}_{\{1<|x|\leq \sigma^{-1}e^{as}\}} xU(\mathrm{d}x)\mathrm{d}s, \\ \upsilon_t &= \tau\frac{\sigma^2}{2a}\left(1 - e^{-2at}\right), \\ V_t(\mathrm{d}x) &= \int_0^t U(\sigma^{-1}e^{as}\mathrm{d}x)\mathrm{d}s. \end{aligned}$$

Proof. See Sato (1999). The proof can be done in the same way as the proof of the Proposition 4.12. □

Corollary 4.15. *The OU type process $(X_t : t \geq 0)$ is almost surely positive if and only if $X_0 \geq 0$ and the BDLP is a subordinator.*

Proof. From the Lévy-Itô decomposition of the process X_t we obtain

$$\mathrm{d}X_t = \left(c_t - \int_{|x|<1} xV_t(\mathrm{d}x)\right)\mathrm{d}t + \sqrt{\upsilon_t}\mathrm{d}W_t + \int_{\mathbb{R}} xJ(\mathrm{d}x \times \mathrm{d}t).$$

The process X is almost surely positive if and only if, for every t

(1) $X_0 \geq 0$

(2) $c_t - \int_{|x|<1} xV_t(\mathrm{d}x) \geq 0$

(3) $\upsilon_t = 0$

(4) the intensity measure V_t of the random meaure J has a support on $\mathbb{R}^+$.

It follows from the previous Proposition 4.14 that the condition (2) is equivalent to the condition

$$b_0 = b - \int_{|x|<1} xU(\mathrm{d}x) \geq 0,$$

that means the drift of L_t is non-negative. The condition (3) is equivalent to $\tau = 0$. The last condition is equivalent to U having support on $\mathbb{R}^+$. From all this we deduce that X_t is almost surely positive if and only if $X_0 \geq 0$ and the BDLP is a subordinator.

□

4.3 GIG–OU type processes

We are in particular interested in the OU type processes associated with GIG distributions. In the Example 4.4 we have shown that GIG distribution is selfdecomposable. Hence from the Proposition 4.9 follows that there exists a stationary OU type process $(X_t : t \geq 0)$ such that $X_t \sim \mathrm{GIG}(\lambda, \delta, \gamma)$. We will call such a process a GIG–OU type process. These processes are positive and thus, from the Corollary 4.15, the background driving Lévy process is a subordinator, i.e. a process with only positive increments. This property implies that the OU type process moves up entirely by jumps and then tails off exponentialy, but stays always positive.

In this section we focus on this particular case of GIG–OU type process and sumarize the known results. By application of the previous development we compute the characteristics of the BDLP.

4.3.1 GIG–OU type process

Let $a > 0$, $\sigma > 0$ and consider the OU type process defined by (4.3), i.e.

$$X_t = e^{-at}X_0 + \int_0^t \sigma e^{-a(t-s)}\mathrm{d}L_s.$$

X_t has $\mathrm{GIG}(\lambda, \delta, \gamma)$ distribution described in Subsection 2.4.4, its generating triplet is given by

$$(c, 0, v(x)\mathrm{d}x),$$

where

$$c = \frac{\delta}{\gamma}\frac{K_{\lambda+1}(\delta\gamma)}{K_\lambda(\delta\gamma)} - \int_1^\infty xv(x)\mathrm{d}x,$$
$$v(x) = x^{-1}\exp\left(-\frac{1}{2}\gamma^2 x\right)\left[\int_0^\infty e^{-x\zeta}g_\lambda(2\delta^2\zeta)\mathrm{d}\zeta + \max\{0, \lambda\}\right], \qquad x \in \mathbb{R}^+,$$

where

$$g_\lambda(z) = \left[(\pi^2/2)z\left\{J^2_{|\lambda|}(\sqrt{z}) + Y^2_{|\lambda|}(\sqrt{z})\right\}\right]^{-1},$$

with J_λ and Y_λ Bessel functions (see Appendix A). The moment generating function of GIG distribution is given by

$$\psi_{X_t}(z) = (1 - 2z/\gamma^2)^{\lambda/2}\frac{K_\lambda\left(\delta\gamma\sqrt{1-2z/\gamma^2}\right)}{K_\lambda(\delta\gamma)}, \quad \mathrm{Re}(z) < \frac{\gamma^2}{2}.$$

Log–moment generating function of the BDLP Using the Proposition 4.11 we compute the log–moment generating function of the BDLP.

Lemma 4.16. *The log–moment generating function of the BDLP L is given by*

$$\theta_{L_1}(z) = \frac{\delta}{\gamma}\frac{a\sigma^{-1}z}{\sqrt{1-2\sigma^{-1}z/\gamma^2}}\frac{K_{\lambda-1}\left(\delta\gamma\sqrt{1-2\sigma^{-1}z/\gamma^2}\right)}{K_\lambda\left(\delta\gamma\sqrt{1-2\sigma^{-1}z/\gamma^2}\right)}. \tag{4.20}$$

Proof. We can rewrite the moment generating function of X_t as

$$\begin{aligned}\psi_{X_t}(z) &= (1 - 2z/\gamma^2)^{\lambda/2}\frac{K_\lambda\left(\delta\gamma\sqrt{1-2z/\gamma^2}\right)}{K_\lambda(\delta\gamma)}\\ &= \frac{(\delta\gamma)^{-\lambda}}{K_\lambda(\delta\gamma)}y^\lambda K_\lambda(y)\Big|_{y=\delta\gamma\sqrt{1-2z/\gamma^2}}, \qquad \mathrm{Re}(z) < \frac{\gamma^2}{2}.\end{aligned}$$

By application of the Proposition 4.11 we compute the logarithm of moment generating function of the BDLP L

$$\theta_{L_1}(\sigma z) = az\frac{\partial \theta_{X_t}(z)}{\partial z} = az\frac{1}{\psi_{X_t}(z)}\frac{\partial \psi_{X_t}(y)}{\partial y}\frac{\partial y(z)}{\partial z}.$$

The partial derivate of $\psi_{X_t}(y)$ can be computed using the recurrence formula (A.6). After computation we obtain the following result.

$$\theta_{L_1}(\sigma z) = \frac{\delta}{\gamma}\frac{az}{\sqrt{1-2z/\gamma^2}}\frac{K_{\lambda-1}(\delta\gamma\sqrt{1-2z/\gamma^2})}{K_\lambda(\delta\gamma\sqrt{1-2z/\gamma^2})}, \qquad \mathrm{Re}(z) < \frac{\gamma^2}{2}.$$

Simple change of variable leads to the statement of the lemma. □

Generating triplet of the BDLP First we show that without loss of generality we may assume that $\sigma = 1$.

Lemma 4.17. *Define $Y_t = \sigma^{-1}X_t$ for every $t \geq 0$. Then Y_t is GIG$(\lambda, \bar{\delta}, \bar{\gamma})$–OU type process, where*

$$\bar{\delta} = \delta\sigma^{-1/2} \qquad \textit{and} \qquad \bar{\gamma} = \gamma\sigma^{1/2},$$

and Y_t solves the following SDE

$$Y_t = e^{-at}Y_0 + \int_0^t e^{-a(t-s)}\mathrm{d}L_s. \tag{4.21}$$

The BDLP of Y_t and X_t is the same.

Proof. Assume that X_t is GIG$(\lambda, \delta, \gamma)$–OU type process as defined in (4.3).
Step 1: We show, that the distribution of Y_t is GIG$(\lambda, \bar{\delta}, \bar{\gamma})$. The probability density function of X_t is given by (2.14). By the theorem about the transformation of random variables we see that the probability density function of Y_t is given by

$$\begin{aligned} g_{Y_t}(y) &= f_{X_t}(\sigma y)\sigma = \\ &= \sigma\left(\frac{\gamma}{\delta}\right)^\lambda \frac{1}{K_\lambda(\delta\gamma)}\sigma^{\lambda-1}y^{\lambda-1}\exp\left\{-\frac{1}{2}\left(\delta^2\sigma^{-1}y^{-1} + \gamma^2\sigma y\right)\right\} = \\ &= \left(\frac{\sigma^{1/2}\gamma}{\sigma^{-1/2}\delta}\right)^\lambda \frac{1}{K_\lambda(\sigma^{-1/2}\delta\sigma^{1/2}\gamma)}y^{\lambda-1}\exp\left\{-\frac{1}{2}\left((\delta\sigma^{-1/2})^2y^{-1} + (\gamma\sigma^{1/2})^2y\right)\right\}. \end{aligned}$$

This is a probability density function of GIG$(\lambda, \bar{\delta}, \bar{\gamma})$.
Step 2: Y_t solves (4.21) follows immediately, since

$$Y_t = \sigma^{-1}\left(e^{-at}X_0 + \int_0^t \sigma e^{-a(t-s)\mathrm{d}L_s}\right) = e^{-at}Y_0 + \int_0^t e^{-a(t-s)}\mathrm{d}L_s.$$

Step 3: The BDLP of Y_t and X_t is the same. We compute the log moment generating function of the BDLP corresponding to X_t and Y_t using the Proposition 4.11. We have from the lemma 4.16

$$\begin{aligned} \theta_{L_1}(z) &= \frac{\delta}{\gamma}\frac{a\sigma^{-1}z}{\sqrt{1-2\sigma^{-1}z/\gamma^2}}\frac{K_{\lambda-1}\left(\delta\gamma\sqrt{1-2\sigma^{-1}z/\gamma^2}\right)}{K_\lambda\left(\delta\gamma\sqrt{1-2\sigma^{-1}z/\gamma^2}\right)} = \\ &= \frac{\bar{\delta}}{\bar{\gamma}}\frac{az}{\sqrt{1-2z/\bar{\gamma}^2}}\frac{K_{\lambda-1}\left(\bar{\delta}\bar{\gamma}\sqrt{1-2z/\bar{\gamma}^2}\right)}{K_\lambda\left(\bar{\delta}\bar{\gamma}\sqrt{1-2z/\bar{\gamma}^2}\right)}, \end{aligned}$$

with $\bar{\delta}$ and $\bar{\gamma}$ given above. Hence the BDLP fo X_t and Y_t has the same moment generating function and thus is the same. □

Lemma 4.18. *The generating triplet of the BDLP L is equal to*

$$(b, 0, u(x)\mathrm{d}x),$$

where

$$b = a\left[\frac{\bar{\delta}}{\bar{\gamma}}\frac{K_{\lambda+1}(\bar{\delta}\bar{\gamma})}{K_\lambda(\bar{\delta}\bar{\gamma})} - v(1) - \int_1^\infty v(x)\mathrm{d}x\right],$$
$$u(x) = a\bar{\delta}^2 \exp\left(-\frac{1}{2}\bar{\gamma}^2 x\right)\int_0^\infty \zeta e^{-x\zeta} g_\lambda(2\delta^2\zeta)\mathrm{d}\zeta + \frac{1}{2}a\bar{\gamma}^2 x v(x).$$

Proof. Follows easily from the Proposition 4.12, Corollary 4.13 and the form of the generating triplet of Y_t. □

4.3.2 IG–OU type process

In the special case of IG–OU type, the process X_t has IG(δ,γ) distribution described in Subsection 2.4.5, its generating triplet is given by

$$(c, 0, v(x)\mathrm{d}x),$$

where

$$c = \frac{\delta}{\gamma}(2\Phi(\gamma) - 1),$$
$$v(x) = \frac{\delta}{\sqrt{2\pi}}x^{-3/2}\exp\left(-\frac{1}{2}\gamma^2 x\right), \qquad x \in \mathbb{R}^+.$$

The log–moment generating function has the form

$$\theta_{X_t}(z) = \delta\gamma\left(1 - \sqrt{1 - 2z/\gamma^2}\right), \qquad \mathrm{Re}(z) < \frac{\gamma^2}{2}.$$

Log–moment generating function of the BDLP As a corollary of the Lemma 4.16, the log–moment generating function of the BDLP in the IG–OU case is given in the following way.

Lemma 4.19. *The log–moment generating function of the BDLP L is given by*

$$\theta_{L_1}(z) = \frac{\bar{\delta}}{\bar{\gamma}}\frac{az}{\sqrt{1 - 2z/\bar{\gamma}^2}}, \tag{4.22}$$

where $\bar{\delta} = \delta\sigma^{-1/2}$ and $\bar{\gamma} = \gamma\sigma^{1/2}$.

Generating triplet of the BDLP The generating triplet of the BDLP in the IG–OU case is given in the following way.

Lemma 4.20. *The generating triplet of the BDLP L is equal to*

$$(b, 0, u(x)\mathrm{d}x),$$

where

$$\begin{aligned} b &= a\bar{\delta}\left(\frac{1}{\bar{\gamma}}(2\Phi(\bar{\gamma})-1)+\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\bar{\gamma}^2\right)\right), \\ u(x) &= a\frac{\bar{\delta}}{2\sqrt{2\pi}}\left(x^{-1}+\bar{\gamma}^2\right)x^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\bar{\gamma}^2 x\right), \qquad x>0. \end{aligned}$$

Theorems about distribution of the IG–OU type process and its BDLP Assume for now that $\sigma = 1$ (we can go back to the case with $\sigma \neq 1$ by taking $\bar{\delta}$ and $\bar{\gamma}$ instead of δ and γ). Barndorff-Nielsen (1998) showed the following result specifying the distribution of the BDLP.

Proposition 4.21. *The BDLP L driving the IG–OU type process $(X_t : t \geq 0)$ with $X_t \sim IG(\delta, \gamma)$ is a sum of two independent Lévy processes, $L_t = L_t^{(1)} + L_t^{(2)}$, where $L_t^{(1)}$ is an IG process with parameters $\delta/2$ and γ, while $L^{(2)}$ is a compound Poisson process of the form*

$$L_t^{(2)} = \gamma^{-1}\sum_{i=1}^{N_t} U_i^2$$

with N_t a Poisson process with parameter $\delta\gamma/2$ and the U_i being independent standard normal and independent of the process N_t.

In the following lemma we derive the density function of the compound Poisson process $L_t^{(2)}$.

Lemma 4.22. *The Lévy process $L^{(2)}$ from the previous Proposition 4.21 is a Poisson–Gamma process,*

$$L_t^{(2)} = X_1 + \cdots + X_{N_t},$$

where N_t is a Poisson process with parameter $\delta\gamma/2$ and X_i are $\Gamma\left(\frac{1}{2},\frac{1}{2}\gamma^2\right)$ distributed. Its density is given by

$$f_2(x) = \exp\left(-\frac{1}{2}\delta\gamma\right)\delta_0(x) + \sum_{k=1}^{\infty}\frac{\left(\frac{1}{2}\delta\gamma\right)^k\left(\frac{1}{2}\gamma^2\right)^{k/2}}{k!\,\Gamma\left(\frac{k}{2}\right)}x^{\frac{k}{2}-1}\exp\left(-\frac{1}{2}\delta\gamma\right)\exp\left(-\frac{1}{2}\gamma^2 x\right). \tag{4.23}$$

Proof. Step 1: Application of the Theorem about random variables transformation (see for example (Anděl, 2005, Theorem 3.5)). Let $Y = \gamma^{-1}U^2$, where $U \sim N(0,1)$. Then the density of Y is given by

$$f_Y(x) = 2\varphi\left(\gamma\sqrt{x}\right)\frac{\gamma}{2\sqrt{x}} = \left(\frac{1}{2}\gamma^2\right)^{\frac{1}{2}}\frac{1}{\sqrt{\pi}}x^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\gamma^2 x\right).$$

Since $\sqrt{\pi} = \Gamma(1/2)$, it follows that Y has $\Gamma\left(\frac{1}{2}, \frac{1}{2}\gamma^2\right)$ distribution.

Step 2: Denote $S_k = X_1 + \cdots + X_k$, where X_i are i.i.d. $\Gamma\left(\frac{1}{2}, \frac{1}{2}\gamma^2\right)$ random variables. Then S_k is $\Gamma\left(\frac{k}{2}, \frac{1}{2}\gamma^2\right)$ distributed random variable. We have

$$\begin{aligned}\mathbb{P}(L_1^{(2)} \le x) &= \sum_{k=1}^{\infty} \mathbb{P}\left(\sum_{i=1}^{N_1} X_i \le x | N_1 = k\right) \mathbb{P}(N_1 = k) = \\ &= \sum_{k=1} \mathbb{P}\left(S_k \le x\right) \mathbb{P}(N_1 = k).\end{aligned}$$

By differentiation w.r.t. x we obtain the density in the form (4.23). □

Zhang and Zhang (2008) also proved the following result specifying the distribution of the stochastic integral

$$I_t = \int_0^t e^{-a(t-s)} \mathrm{d}L_s$$

that appears in the definition of the OU–type process.

Proposition 4.23. *Assume that $\gamma > 0$. Then the random variable I_t is a sum of two independent Lévy processes, $I_t = I_t^{(1)} + I_t^{(2)}$, where $I_t^{(1)}$ is an IG process with parameters $\delta\left(1 - e^{-\frac{1}{2}at}\right)$ and γ, while $I_t^{(2)}$ is a compound Poisson process of the form*

$$I_t^{(2)} = \sum_{i=1}^{N_t} W_i^t$$

with N_t a Poisson process with intensity $\delta\gamma\left(1 - e^{-\frac{1}{2}at}\right)$ and W_i^t being independent random variables independent of N_t having a common density function

$$f_t(x) = \begin{cases} \frac{\gamma^{-1}}{\sqrt{2\pi}} x^{-3/2} \left(e^{\frac{1}{2}at} - 1\right)^{-1} \left(e^{-\frac{1}{2}\gamma^2 x} - e^{-\frac{1}{2}\gamma^2 x e^{at}}\right), & x > 0 \\ 0, & \textit{otherwise}. \end{cases} \tag{4.24}$$

Proof. See Zhang and Zhang (2008). □

4.3.3 Gamma–OU type process

In the special case of Gamma–OU type, the process X_t has Gamma(λ, γ) distribution described in Subsection 2.4.6, its generating triplet is given by

$$(c, 0, v(x)\mathrm{d}x),$$

where

$$\begin{aligned} c &= \frac{\lambda}{\beta}\left(1 - \exp(-\beta)\right), \\ v(x) &= \lambda x^{-1} \exp\left(-\beta x\right), \qquad x \in \mathbb{R}^+, \quad \beta = \gamma^2/2. \end{aligned}$$

The log–moment generating function has the form

$$\theta_{X_t}(z) = -\lambda \ln\left(1 - z/\beta\right), \qquad \mathrm{z} \ne \beta.$$

Log–moment generating function of the BDLP By application of the Proposition 4.11, the log–moment generating function of the BDLP in the Gamma–OU case is given in the following way.

Lemma 4.24. *The log–moment generating function of the BDLP L is given by*

$$\theta_{L_1}(z) = a\lambda\tilde{\beta}\frac{z}{1 - z/\tilde{\gamma}}, \qquad Re(z) \neq \tilde{\beta}, \tag{4.25}$$

where $\tilde{\beta} = \beta\sigma$.

Generating triplet of the BDLP The generating triplet of the BDLP in the Gamma–OU case is given in the following way.

Lemma 4.25. *The generating triplet of the BDLP L is equal to*

$$(b, 0, u(x)\mathrm{d}x),$$

where

$$\begin{aligned} b &= a\lambda\tilde{\beta}^{-1}\left(1 - e^{-\tilde{\beta}} - \tilde{\beta}e^{-\tilde{\beta}}\right), \\ u(x) &= a\lambda\tilde{\beta}\exp(-\tilde{\beta}x), \qquad x > 0. \end{aligned}$$

Chapter 5

Term structure models

5.1 Motivation

Our approach takes its motivation from a most popular and widely accepted class of short rate models in continuous time. The short rate is modelled as a one-dimensional diffusion process and we focus on the case where the diffusion is a mean-reverting Ornstein–Uhlenbeck process, that is usually termed the "Vašíček model" in the literature. Our aim is to find a generalization of this model for Lévy drivers by considering the short rate to be a mean-reverting Ornstein–Uhlenbeck type process that was studied in Chapter 4.

Consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with a filtration $\mathbb{F} = (\mathcal{F}_t : t \geq 0)$ satisfying the usual conditions and let $\mathbb{Q}$ be a martingale measure such that W_t^* is $(\mathbb{Q}, \mathbb{F})$-Brownian motion. The short rate dynamics is modelled through a SDE with

$$\mathrm{d}r_t = (\vartheta(t) - a(t)r_t)\mathrm{d}t + \sigma(t)r_t^\gamma \mathrm{d}W_t^*, \tag{5.1}$$

where ϑ, a, $\sigma : \mathbb{R}^+ \to \mathbb{R}$ are deterministic and locally bounded functions and $0 \leq \gamma \leq 1$ is a constant. This model goes at least to Hull and White (1990). Special cases of this model are:

Ho-Lee model If we set $\gamma = 0$, $a \equiv 0$ and $\sigma(t) = \sigma$, we obtain the continuous version of the Ho-Lee model introduced in discrete time in Ho and Lee (1986).

$$\mathrm{d}r_t = \vartheta(t)\mathrm{d}t + \sigma \mathrm{d}W_t^*. \tag{5.2}$$

Drawbacks of this model are that it incorporates no mean reversion and the short rate can be negative.

Vašíček model By setting $\gamma = 0$ we obtain the extended Vašíček model in which the short rate process is mean reverting and we can find explicit formulas for bond prices. One drawback of assuming $\gamma = 0$ is that the short rate can become negative.

$$\mathrm{d}r_t = (\vartheta(t) - a(t)r_t)\mathrm{d}t + \sigma(t)\mathrm{d}W_t^*.$$

The model introduced by Vašíček (1977) consider time-independent coefficients and hence reduces to

$$\mathrm{d}r_t = (\vartheta - ar_t)\mathrm{d}t + \sigma \mathrm{d}W_t^*. \tag{5.3}$$

Cox, Ingerssol, Ross (CIR) model In this model we set $\gamma = 1/2$ and we obtain a mean reverting short rate process that is moreover always positive (can become zero under some circumstances).

$$\mathrm{d}r_t = (\vartheta(t) - a(t)r_t)\mathrm{d}t + \sigma(t)\sqrt{r_t}\mathrm{d}W_t^*.$$

The CIR model developed in Cox et al. (1985) consider constant parameters and reduces to

$$\mathrm{d}r_t = (\vartheta - ar_t)\mathrm{d}t + \sigma\sqrt{r_t}\mathrm{d}W_t^*.$$

These kinds of interest rate models specifying the short rate process dynamics are quite popular by the practitioners because they are tractable. However, there is a deeper theory behind these models. All these models are special cases of the *HJM (Heath-Jarrow-Morton) model*, introduced in Heath et al. (1992). In the HJM framework we describe the whole term structure of interest rates by modelling the instantaneous forward rates $f(t,T)$ as a diffusion processes

$$\mathrm{d}f(t,T) = \alpha(t,T)\mathrm{d}t + \sigma(t,T)\mathrm{d}W_t, \tag{5.4}$$

and we derive arbitrage-free framework for the term structure by way of specifying the drift term α of the diffusion in terms of the volatility σ.

In the present Chapter we generalize the HJM model into the Lévy processes framework by way of replacing the Brownian motion in (5.4) by a general Lévy process. We proceed in three steps. As a first step we determine in Section 5.2 the no-aritrage dynamics of the discounted bond prices as induced by an "affine linear" Lévy dynamics of the forward rate, and thus obtain a risk-neutral dynamics of the short rate as a consequence. We have thus a machine for constructiong no-arbitrage Lévy driven short rates. As a second step we formulate in Section 5.3 criteria for thus constructed processes to meet the "stylized facts" required for short rates: in particular "positivity" and "convergence to a mean rate". Non-emptiness of the theory is provided as a third step: in Section 5.4 by demonstrating the case of OU type processes associated with the generalized inverse Gaussian (GIG) and normal inverse Gaussian (NIG) distributions the upshot here is an explicitely given short rate process, which directly generalizes the Vašíček model (5.3), but which now, in the GIG case, stays positive.

5.2 Presentation of the model, construction of its no-arbitrage dynamics

The term structure of interest rates is expressed by the following two families of securities.

(1) The *savings account* $B = (B_t : t \geq 0)$. For any time $t \geq 0$, B_t denotes the time t value of 1 unit of currency continuously compounded starting at time 0.

(2) The *zero coupon bonds* $\big(P(t,T) : 0 \leq t \leq T\big)$ with maturities in $[0,T^*]$. For any $T \in [0,T^*]$, $P(t,T)$ denotes the time t value of the zero coupon bond which pays 1 unit of currency at maturity T.

In particular, $B_0 = 1$ and $P(T,T) = 1$ in the equilibrium. Assuming differentiability with respect to time, we have the infinitesimal characterisations (with $0 \leq t \leq T \leq T^*$):

(3) The *short rate*

$$r_t = \frac{\partial \log B_t}{\partial t}$$

is the time t instantaneous rate for continuously compounded risk-free borrowing and lending over the infinitesimally short time period $[t, t + \mathrm{d}t)$.

(4) The *forward rate*

$$f(t,T) = -\frac{\partial \log P(t,T)}{\partial T}$$

is the time t instantaneous rate for continuously compounded risk-free borrowing and lending over the infinitesimally short time period $[T, T + \mathrm{d}T)$.

We may also equivalently write

$$B_t = \exp\left(\int_0^t r_u \mathrm{d}u\right) \qquad \text{and} \qquad P(t,T) = \exp\left(-\int_t^T f(t,u)\mathrm{d}u\right).$$

We consider $t \mapsto B_t$ and $t \mapsto P(t,T)$ for any $T \in [0, T^*]$ to be sufficiently nice process on filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We are looking for a probability measure $\mathbb{Q} \sim \mathbb{P}$ such that the discounted bond prices process

$$\left(\tilde{P}(t,T) \stackrel{\text{def}}{=} \frac{P(t,T)}{B_t}, 0 \le t \le T\right),$$

for any $T \in [0, T^*]$ is a $(\mathbb{Q}, \mathbb{F})$-martingale. Such measure is called the martingale measure. The no-arbitrage dynamics of the bond price process will be developed through the no-arbitrage dynamics of the instantaneous forward rate process $t \mapsto f(t,T)$, for all $T \in [0, T^*]$.

5.2.1 Modelling of no-arbitrage dynamics of forward rate

Let $(L_t : t \ge 0)$ be a Lévy process on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. We assume that the instantaneous forward rate follows, for any $T \in [0, T^*]$, a process of the form

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)\mathrm{d}s + \int_0^t \sigma(s,T)\mathrm{d}L_s, \qquad t \in [0,T], \tag{5.5}$$

where $T \mapsto f(0,T)$ is the initial forward rate structure and $\alpha, \sigma : [0, T^*]^2 \to \mathbb{R}$ are deterministic coefficients.

Remark 5.1. The initial forward rate structure can be obtained from a current yield curve in the following way. Denote be $y(T)$ the value of the interest rate for time period $[0,T)$, i.e. the value of the yield curve for maturity T. Then the following relationship between $f(0,T)$ and $y(T)$ holds:

$$f(0,T) = y(T) + T\frac{\partial y(T)}{\partial T}.$$

We will consider the following four assumptions on the process $f(t,T)$:

(MG) The domain of definition of the moment generating function of L_1 contains a nonempty open neigbourhood of a translate of the imaginary axis.

(Existence) For any $T \in [0, T^*]$ the process $(f(t,T) : t \in [0,T])$ is well defined, càdlàg semimartingale.

(Fubini 1) For any $0 \leq t \leq T \leq T^*$ we have

$$\begin{aligned}\int_0^T \int_0^{u\wedge t} \alpha(s,u)\mathrm{d}s\,\mathrm{d}u &\stackrel{\text{def}}{=} \int_0^T \int_0^T H(\alpha(s,u))\mathrm{d}s\,\mathrm{d}u = \\ &= \int_0^T \int_0^T H(\alpha(s,u))\mathrm{d}u\,\mathrm{d}s = \int_0^t A(s,T)\mathrm{d}s < \infty\end{aligned}$$

(Fubini 2) For any $0 \leq t \leq T \leq T^*$ we have

$$\begin{aligned}\int_0^T \int_0^{u\wedge t} \sigma(s,u)\mathrm{d}L_s\,\mathrm{d}u &\stackrel{\text{def}}{=} \int_0^T \int_0^T H(\sigma(s,u))\mathrm{d}L_s\,\mathrm{d}u = \\ &= \int_0^T \int_0^T H(\sigma(s,u))\mathrm{d}u\,\mathrm{d}L_s = \int_0^t \Sigma(s,T)\mathrm{d}L_s < \infty\end{aligned}$$

Here we adopt the notation:

$$H(f(s,u)) = \mathbf{1}_{X_{t,T}}(s,u)f(s,u), \qquad (s,u) \in [0,T^*]^2,$$

for any function $f : [0,T^*]^2 \to \mathbb{R}$, with the definition

$$\begin{aligned}X_{t,T} &= \{(u,s) \in [0,T^*]^2 | 0 \leq s \leq t, s \leq u \leq T\} = \\ &= \{(u,s) \in [0,T^*]^2 | 0 \leq u \leq T, 0 \leq s \leq u \wedge t\},\end{aligned}$$

and we put

$$A(s,T) = \int_{s\wedge T}^T \alpha(s,u)\mathrm{d}u, \qquad \Sigma(s,T) = \int_{s\wedge T}^T \sigma(s,u)\mathrm{d}u.$$

The condition (Existence) assures that the process given by (5.5) is well-defined, i.e. that the deterministic and stochastic integrals are well-defined. The condition (Fubini 1) is the condition of Fubini theorem, under which it is possible to change the order of integration. The third condition (Fubini 2) is the Fubini theorem in stochastic version. What regards sufficient conditions for (MG), (Existence), (Fubini 1) and (Fubini 2) to hold we consider the following 5 conditions:

(C0) The set $\{z : |\psi_{L_1}(z)| < \infty\}$ contains a subset of the complex plane $(-a,b) \times \mathrm{i}\mathbb{R}$ for real $a, b > 0$, where ψ_{L_1} denotes the moment generating function of L_1.

(C1) $\alpha \in L^1([0,T^*],\mathbb{R})$.

(C2) $\sigma \in L^1_{\mathrm{bd}}([0,T^*],\mathbb{R})$.

(C3) $(T \mapsto f(0,T)) \in C([0,T^*],\mathbb{R})$.

For financial relevance they are to be supplemented by

(C4) $\sigma(\{(s,u) \in [0,T^*]^2 : s > u\}) = \{0\}$ (vanishing of the σ "above the diagonal").

In fact we have the following results, the first one being obvious.

Fact 5.2. *Assume that (C0) holds. Then the condition (MG) is satisfied.*

Proposition 5.3. *Assume that (C1), (C2) holds. Then the condition (Existence) is satisfied.*

Proof. The existence of the deterministic integral is immediate; since $\alpha \in L^1([0,T^*],\mathbb{R})$ we have

$$\int_0^t \alpha(s,T)\mathrm{d}s < \infty, \qquad \text{for all } t \in [0,T].$$

The existence of the stochastic integral follows from the Theorem 3.6: L is a semimartingale and σ is bounded and deterministic function, hence predictable. □

Proposition 5.4. *We have the following implications: (C1) implies (Fubini 1) and (C2) implies (Fubini 2).*

This result relies on the stochastic Fubini's theorem in the following form which combines (Protter, 2004, Chapter IV, Theorem 63) and (Protter, 2004, Chapter IV, Theorem 64).

Theorem 5.5 (Stochastic version of the Fubini's theorem)**.** *Let X be a $(\mathbb{F},\mathbb{P})$-semimartingale, $(A,\mathcal{A})$ be a measurable space; $H_T^a \overset{\text{def}}{=} H(a,T,\omega)$ be a bounded $\mathcal{A}$-predictably meausurable function, μ be a finite measure on $\mathcal{A}$. Then, on choosing an appropriate version if necessary,*

$$Z_T^a \overset{\text{def}}{=} \int_0^T H_s^a \mathrm{d}X_s$$

can be regarded as an $\mathcal{A} \otimes \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{F}$ meausurable function such that $(Z_T^a, T \geq 0)$, for each $a \in A$, is a càdlàg and adapted process.

Moreover, also on choosing an appropriate càdlàg version if necessary, we have:

$$\begin{aligned}\int_A \left(\int_0^T H_s^a \mathrm{d}X_s\right)\mu(\mathrm{d}a) &\overset{\text{def}}{=} \int_A Z_T^a \mu(\mathrm{d}a)\\ &= \int_0^T \left(\int_A H_s^a \mu(\mathrm{d}a)\right)\mathrm{d}X_s,\end{aligned}$$

for any $T \geq 0$.

Proof of Proposition 5.4. The first implication is the standard Fubini's theorem that can be found for example in (Rudin, 1970, Chapter 7, Theorem 7.8).

The second implication is an application of the Theorem 5.5. The process $H_s^u(\omega) = H(\sigma(s,u))(\omega)$, for all $s,u \in [0,T]$, is bounded as follows from the boundedness of σ; the measure μ is in our case the Lebesgue measure hence finite on $\mathbb{R}$ and $X = L$. Then, with $A = [0,T]$, we have

$$\begin{aligned}\int_0^T \left(\int_0^T H(\sigma(s,u))\mathrm{d}L_s\right)\mathrm{d}u &= \int_A \left(\int_0^T H_s^u \mathrm{d}X_s\right)\mathrm{d}u\\ &= \int_0^T \left(\int_A H_s^u \mathrm{d}u\right)\mathrm{d}X_s = \int_0^T \left(\int_0^T H(\sigma(s,u))\mathrm{d}u\right)\mathrm{d}L_s,\end{aligned}$$

what is exactly the (Fubini 2). □

5.2.2 Implied bond price dynamics under the statistical measure

Proposition 5.6. *Assume that (MG), (Existence), (Fubini 1) and (Fubini 2) hold. Then the bond price is given by*

$$P(t,T) = P(0,T)\ B_t\ \exp\left(-\int_0^t A(s,T)\mathrm{d}s - \int_0^t \Sigma(s,T)\mathrm{d}L_s\right), \tag{5.6}$$

for any $T \in [0,T^*]$ and any $t \in [0,T]$, recalling the definitions

$$A(s,T) = \int_s^T \alpha(s,u)\mathrm{d}u \qquad \textit{and} \qquad \Sigma(s,T) = \int_s^T \sigma(s,u)\mathrm{d}u,$$

for any $0 \le s \le T$.

Proof. Recall that

$$P(t,T) = \exp\left(-\int_t^T f(t,u)\mathrm{d}u\right), \quad B_t = \exp\left(\int_0^t r_u\mathrm{d}u\right) \quad \text{and } r_u = f(u,u).$$

As a first step, we look at the log of the factor $P(0,T)B_t$ of the RHS of (5.6). Reverting to rates,

$$\log(P(0,T)\ B_t) = -\int_0^T f(0,u)\mathrm{d}u + \int_0^t r_u\mathrm{d}u.$$

Express r_u by the SDE definition of forward rates using $r_u = f(u,u)$ to obtain the representation

$$\log(P(0,T)\ B_t) = -\int_0^T f(0,u)\mathrm{d}u + \int_0^t \left(f(0,u) + \int_0^u \alpha(s,u)\mathrm{d}s + \int_0^u \sigma(s,u)\mathrm{d}L_s\right)\mathrm{d}u.$$

On cancelation of the integrals with respect to $f(0,u)$ we hence arrive at

$$\log(P(0,T)\ B_t) = -\int_t^T f(0,u)\mathrm{d}u + \int_0^t\int_0^u \alpha(s,u)\mathrm{d}s\mathrm{d}u + \int_0^t\int_0^u \sigma(s,u)\mathrm{d}L_s\mathrm{d}u.$$

Used in the full expression of the RHS of (5.6) this gives:

$$\begin{aligned}\log(RHS) &= \log(P(0,T)\ B_t) - \int_0^t A(s,T)\mathrm{d}s - \int_0^t \Sigma(s,T)\mathrm{d}L_s \\ &= -\int_t^T f(0,u)\mathrm{d}u + \Sigma_1 + \Sigma_2,\end{aligned}$$

where

$$\begin{aligned}\Sigma_1 &= \int_0^t\int_0^u \alpha(s,u)\mathrm{d}s\ \mathrm{d}u - \int_0^t A(s,T)\mathrm{d}s, \\ \Sigma_2 &= \int_0^t\int_0^u \sigma(s,u)\mathrm{d}L_s\ \mathrm{d}u - \int_0^t \Sigma(s,T)\mathrm{d}L_s.\end{aligned}$$

To the summands Σ_1 and Σ_2 apply the respective (Fubini 1) and (Fubini 2), whence

$$\begin{aligned}\Sigma_1 &= -\int_t^T\int_0^t \alpha(s,u)\mathrm{d}s\ \mathrm{d}u, \\ \Sigma_2 &= -\int_t^T\int_0^t \sigma(s,u)\mathrm{d}L_s\ \mathrm{d}u.\end{aligned}$$

As a consequence

$$\log(RHS) = -\int_t^T f(0,u)\,\mathrm{d}u - \int_t^T \int_0^t \alpha(s,u)\,\mathrm{d}s\,\mathrm{d}u - \int_t^T \int_0^t \sigma(s,u)\,\mathrm{d}L_s\,\mathrm{d}u.$$

Referring to the SDE defining the rates $f(t,u)$, namely

$$f(t,u) = f(0,u) + \int_0^t \alpha(s,u)\mathrm{d}s + \int_0^t \sigma(s,u)\mathrm{d}L_s, \qquad \text{for } t \leq u,$$

we thus obtain on inspection:

$$\log(RHS) = -\int_t^T f(t,u)\mathrm{d}u,$$

which is the log of the LHS of (5.6), as was to be shown. □

The strategy of obtaining dynamics in the risk neutral setting is as follows. We keep our statistical measure as the risk neutral one; we define $\mathbb{Q} \stackrel{\text{def}}{=} \mathbb{P}$. Then, for any $T \in [0,T^*]$ we make the discounted bond price process a $(\mathbb{P},\mathbb{F})$–martingale on $[0,T]$ by adapting the drift coefficients α, by the way of specifying the integrated coefficients $A(s,T)$.

5.2.3 No-arbitrage dynamics of the bond prices

We obtain a no-arbitrage dynamics of bond prices that render the statistical measure $\mathbb{P}$ a martingale measure by choosing the drift appropriately as follows.

Proposition 5.7. *Assume that (MG), (Existence), (Fubini 1) and (Fubini 2) hold. For any $T \in [0,T^*]$, with the definition*

$$A(s,T) = \theta_{L_1}(-\Sigma(s,T)), \qquad s \in [0,T],$$

the discounted bond price price process

$$\left(\tilde{P}(t,T),\, t \in [0,T]\right), \qquad \textit{with} \qquad \tilde{P}(t,T) = \frac{P(t,T)}{B_t}$$

is a $(\mathbb{P},\mathbb{F})$–martingale.

Proof. We look at the process given by a stochastic integral

$$X_t = \int_0^t (-\Sigma(s,T))\mathrm{d}L_s.$$

Step 1: By application of the Key theorem 3.9 we obtain:

$$\mathrm{E}[\exp(X_t)] = \exp\left(\int_0^t \theta_{L_1}(-\Sigma(s,T))\mathrm{d}s\right) = \exp\left(\int_0^t A(s,T)\mathrm{d}s\right).$$

Step 2: In the bond price formula (5.6) this yields

$$\tilde{P}(t,T) \stackrel{\text{def}}{=} \frac{P(t,T)}{B_t} = P(0,T)\frac{\exp(X_t)}{\mathrm{E}[\exp(X_t)]}.$$

We want to show that $(\tilde{P}(t,T) : t \in [0,T])$ is a $(\mathbb{P},\mathbb{F})$–martingale for every $T \in [0,T^*]$, that means by definition that

$$\mathrm{E}\big[\tilde{P}(t,T)|\mathcal{F}_s\big] = \tilde{P}(s,T), \qquad \text{for all } s \le t.$$

This is equivalent to showing that

$$\frac{\mathrm{E}[\exp(X_t)|\mathcal{F}_s]}{\mathrm{E}[\exp(X_t)]} = \frac{\exp(X_s)}{\mathrm{E}[\exp(X_s)]}, \qquad \text{for all } s \le t.$$

Step 3: It follows from the Theorem 3.10 that X_t is a process with independent increments. Hence $X_t - X_s$ is independent of X_s and $X_t - X_s$ is also independent of $\mathcal{F}_s$. One thus has

$$\begin{aligned}
\mathrm{E}[\exp(X_t - X_s)\exp(X_s)|\mathcal{F}_s] &= \mathrm{E}[\exp(X_t - X_s)|\mathcal{F}_s]\mathrm{E}[\exp(X_s)|\mathcal{F}_s], && \text{for all } s \le t,\\
\mathrm{E}[\exp(X_t - X_s)|\mathcal{F}_s] &= \mathrm{E}[\exp(X_t - X_s)], && \text{for all } s \le t,\\
\mathrm{E}[\exp(X_s)|\mathcal{F}_s] &= \exp(X_s), && \text{for all } s,
\end{aligned}$$

because X_s is $\mathcal{F}_s$-measurable. Hence, for all $s \le t$,

$$\begin{aligned}
\mathrm{E}[\exp(X_t)|\mathcal{F}_s] &= \mathrm{E}[\exp(X_t - X_s)\exp(X_s)|\mathcal{F}_s] =\\
&= \mathrm{E}[\exp(X_t - X_s)|\mathcal{F}_s]\ \mathrm{E}[\exp(X_s)|\mathcal{F}_s] =\\
&= \mathrm{E}[\exp(X_t - X_s)]\ \exp(X_s) =\\
&= \frac{\exp(X_s)}{\mathrm{E}[\exp(X_s)]}\mathrm{E}[\exp(X_t)].
\end{aligned}$$

We conclude that the process

$$\left(\frac{\exp(X_t)}{\mathrm{E}[\exp(X_t)]} : t \in [0,T]\right)$$

is a $(\mathbb{P},\mathbb{F})$–martingale and so is $(\tilde{P}(t,T) : t \in [0,T])$. □

The no-arbitrage dynamics of $P(t,T)$ translates into the no-arbitrage dynamics of $f(t,T)$ in the following way.

Corollary 5.8. *Assuming (C0) – (C4), the definition*

$$\alpha(s,t) \stackrel{\text{def}}{=} \frac{\partial}{\partial t}\theta_{L_1}(-\Sigma(s,t))$$

for any $0 \le s \le t \le T^*$ *yields a bond price dynamics for which* $(\tilde{P}(t,T) : t \in [0,T])$ *is a* $(\mathbb{P},F)$*–martingale, for any* $T \in [0,T^*]$.

Proof. Recalling the definition

$$A(s,T) = \int_s^T \alpha(s,u)\mathrm{d}u,$$

we have

$$\alpha(s,u) = \frac{\partial}{\partial u}A(s,u) = \frac{\partial}{\partial u}\theta_{L_1}(-\Sigma(s,u)).$$

□

The short rate process no-arbitrage dynamics is given in the following result.

Corollary 5.9. *Assume (C0) – (C4). Then the following two assertions hold*

(i) The no-arbitrage dynamics of the short rate r is given by

$$r_t = f(0,t) + \int_0^t \frac{\partial}{\partial t}\theta_{L_1}(-\Sigma(s,t))\mathrm{d}s + \int_0^t \sigma(s,t)\mathrm{d}L_s, \qquad t \in [0,T^*]. \tag{5.7}$$

(ii) Assume that moreover the compatibilities

$$\frac{\partial}{\partial T}\Sigma(s,T) = \varepsilon\frac{\partial}{\partial s}\Sigma(s,T) \tag{S_ε}$$

hold for some $\varepsilon \neq 0$, then we have in the drift in (5.7)

$$\int_0^t \frac{\partial}{\partial t}\theta_{L_1}(-\Sigma(s,t))\mathrm{d}s = \varepsilon(\theta_{L_1}(0) - \theta_{L_1}(-\Sigma(0,t))) \tag{5.8}$$

Proof. In view of the compatibilities $r_t = f(t,t)$ for all $t \in [0,T^*]$ the first assertion is immediate from the Corollary 5.8. In the presence of (S_ε) we moreover have

$$\begin{aligned}
\frac{\partial}{\partial t}\theta_{L_1}(-\Sigma(s,t)) &= \theta'_{L_1}(-\Sigma(s,t))\Big(-\frac{\partial}{\partial t}\Sigma(s,t)\Big) = \\
&= \theta'_{L_1}(-\Sigma(s,t))\Big(-\varepsilon\frac{\partial}{\partial s}\Sigma(s,t)\Big) = \\
&= \varepsilon\frac{\partial}{\partial s}\theta_{L_1}(-\Sigma(s,t)),
\end{aligned}$$

whence we obtain (5.8) on integration of the drift term in (5.7). □

In the no-arbitrage setting of the Proposition 5.7 we have the following re-interpretation of the bond price formula as "risk neutral bond price formula".

Corollary 5.10. *In the situation of the Proposition 5.7 and the Corollary 5.9, for any $T \in [0,T^*]$, we have*

$$P(t,T) = \mathrm{E}\left[\exp\left(-\int_t^T r_u \mathrm{d}u\right)|\mathcal{F}_t\right], \tag{5.9}$$

for any $t \in [0,T]$.

Proof. As a *first step* we look at the integral

$$\int_t^T r_u \mathrm{d}u$$

and express r_u as $f(u,u)$, namely

$$r_u = f(u,u) = f(0,u) + \int_0^u \alpha(s,u)\mathrm{d}s + \int_0^u \sigma(s,u)\mathrm{d}L_s,$$

to obtain

$$\int_t^T r_u \mathrm{d}u = \int_t^T f(0,u)\mathrm{d}u + S_1 + S_2, \tag{5.10}$$

where

$$\begin{aligned} S_1 &= \int_t^T \int_0^u \alpha(s,u)\mathrm{d}s\, \mathrm{d}u, \\ S_2 &= \int_t^T \int_0^u \sigma(s,u)\mathrm{d}L_s\, \mathrm{d}u. \end{aligned}$$

Application of (Fubini 1) and (Fubini 2) to S_1 and S_2 respectively yields further decomposition

$$S_1 = S_{1,1} + S_{1,2}, \tag{5.11}$$

with

$$\begin{aligned} S_{1,1} &= \int_0^t \int_t^T \alpha(s,u)\mathrm{d}u\, \mathrm{d}s, \\ S_{1,2} &= \int_t^T \int_s^T \alpha(s,u)\mathrm{d}u\, \mathrm{d}s = \int_t^T A(s,T)\mathrm{d}s, \end{aligned}$$

recalling the definition

$$A(s,t) = \int_s^T \alpha(s,u)\mathrm{d}u$$

for the second indentity in $S_{1,2}$; and

$$S_2 = S_{2,1} + S_{2,2}, \tag{5.12}$$

with

$$\begin{aligned} S_{2,1} &= \int_0^t \int_t^T \sigma(s,u)\mathrm{d}u\, \mathrm{d}L_s, \\ S_{2,2} &= \int_t^T \int_s^T \sigma(s,u)\mathrm{d}u\, \mathrm{d}L_s = \int_t^T \Sigma(s,T)\mathrm{d}s, \end{aligned}$$

recalling the definition

$$\Sigma(s,t) = \int_s^T \sigma(s,u)\mathrm{d}u$$

for the second indentity in $S_{2,2}$.

As a *second step* we study the expression (5.9) of the Corollary in light of the decompositions (5.10), (5.11) and (5.12). The $\mathcal{F}_t$-measurability of the factors $S_{1,1}$, $S_{1,2}$ and $S_{2,1}$ yields

$$\mathrm{E}\left[\exp\left(-\int_t^T r_u \mathrm{d}u\right)|\mathcal{F}_t\right] = \exp\left(-\int_t^T f(0,u)\mathrm{d}u - S_{1,1} - S_{2,1}\right) \times R,$$

where

$$R = \exp(-S_{1,2})\mathrm{E}\left[\exp(-S_{2,2})|\mathcal{F}_t\right].$$

The factor R on the right hand side we now analyse taking into account our risk-neutral choice of coefficients

$$A(s,T) = \theta_{L_1}(-\Sigma(s,T)), \qquad s \leq T,$$

which yields

$$\exp(-S_{1,2}) = \exp\left(-\int_t^T \theta_{L_1}(-\Sigma(s,T))\mathrm{d}s\right).$$

Application of the Key theorem 3.9, on the other hand, gives

$$\begin{aligned}\mathrm{E}\left[\exp(-S_{2,2})|\mathcal{F}_t\right] &= \mathrm{E}\left[\exp\left(-\int_t^T \Sigma(s,T)\mathrm{d}s\right)|\mathcal{F}_t\right] \\ &= \exp\left(\int_t^T \theta_{L_1}(-\Sigma(s,T))\mathrm{d}s\right).\end{aligned}$$

Collecting terms,

$$R = \exp\left(-\int_t^T \theta_{L_1}(-\Sigma(s,T))\mathrm{d}s\right) \times \exp\left(\int_t^T \theta_{L_1}(-\Sigma(s,T))\mathrm{d}s\right) = 1,$$

and hence

$$\mathrm{E}\left[\exp\left(-\int_t^T r_u \mathrm{d}u\right)|\mathcal{F}_t\right] = \exp\left(-\int_t^T f(0,u)\mathrm{d}u - S_{1,1} - S_{2,1}\right)$$

As a *third step* we study how to reverse the application of the (Fubini) in this representation. For this we start by asking about the difference of the $S_{k,1}$ and the integrals from 0 to t of $A(s,T)$ and $\Sigma(s,T)$ respectively, namely

$$\begin{aligned} S_{1,1} &= \int_0^t \int_t^T \alpha(s,u)\mathrm{d}u\ \mathrm{d}s = \int_0^t A(s,T)\mathrm{d}s - T_1 \\ \text{with} \qquad & T_1 = \int_0^t \int_s^t \alpha(s,u)\mathrm{d}u\ \mathrm{d}s,\end{aligned}$$

as well as

$$\begin{aligned} S_{2,1} &= \int_0^t \int_t^T \sigma(s,u)\mathrm{d}u\ \mathrm{d}L_s = \int_0^t \Sigma(s,T)\mathrm{d}s - T_2 \\ \text{with} \qquad & T_2 = \int_0^t \int_s^t \sigma(s,u)\mathrm{d}u\ \mathrm{d}L_s.\end{aligned}$$

Application of the (Fubini 1) and (Fubini 2) shows

$$\begin{aligned} T_1 &= \int_0^t \int_0^u \alpha(s,u)\mathrm{d}s\ \mathrm{d}u, \\ T_2 &= \int_0^t \int_0^u \sigma(s,u)\mathrm{d}L_s\ \mathrm{d}u,\end{aligned}$$

whence the representation

$$\begin{aligned}
\int_0^t f(0,u)\mathrm{d}u + T_1 + T_2 &= \int_0^t \left(f(0,u) + \int_0^u \alpha(s,u)\mathrm{d}s + \int_0^u \sigma(s,u)\mathrm{d}L_s \right) \mathrm{d}u \\
&= \int_0^t f(u,u)\mathrm{d}u \\
&= \int_0^t r_u \mathrm{d}u.
\end{aligned}$$

Returning to the representation:

$$\mathrm{E}\left[\exp\left(-\int_t^T r_u \mathrm{d}u \right) |\mathcal{F}_t \right] = \exp\left(-\int_t^T f(0,u)\mathrm{d}u - S_{1,1} - S_{2,1} \right).$$

First rewrite the exponent of the RHS tautologicaly as

$$\begin{aligned}
&-\int_t^T f(0,u)\mathrm{d}u - S_{1,1} - S_{2,1} \\
&= -\int_0^T f(0,u)\mathrm{d}u + \int_0^t f(0,u)\mathrm{d}u - \left(\int_0^t A(s,T)\mathrm{d}s - T_1 \right) - \left(\int_0^t \Sigma(s,T)\mathrm{d}L_s - T_2 \right) \\
&= -\int_0^T f(0,u)\mathrm{d}u + \left(\int_0^t f(0,u)\mathrm{d}u + T_1 + T_2 \right) - \int_0^t A(s,T)\mathrm{d}s - \int_0^t \Sigma(s,T)\mathrm{d}L_s.
\end{aligned}$$

Recall that the second summand in brackets is the integral of r_u from 0 to t; we hence obtain

$$\begin{aligned}
&\mathrm{E}\left[\exp\left(-\int_t^T r_u \mathrm{d}u \right) |\mathcal{F}_t \right] \\
&= \exp\left(-\int_0^T f(0,u)\mathrm{d}u \right) \exp\left(\int_0^t r_u \mathrm{d}u \right) \exp\left(-\int_0^t A(s,T)\mathrm{d}s - \int_0^t \Sigma(s,T)\mathrm{d}L_s \right) \\
&= P(0,T)\ B_t\ \exp\left(-\int_0^t A(s,T)\mathrm{d}s - \int_0^t \Sigma(s,T)\mathrm{d}L_s \right),
\end{aligned}$$

as was to be shown. □

5.3 Properties of the short rate

Under (C0) – (C4), we consider the short rate process r just constructed:

$$r_t = m_t + \int_0^t \sigma(s,t)\mathrm{d}L_s, \qquad t \in [0,T^*], \tag{5.13}$$

where

$$m_t = f(0,t) + \int_0^t \frac{\partial}{\partial t}\theta_{L_1}(-\Sigma(s,t)).$$

We ask for the compatibility of the construction with the "stylized facts" postulated for a short rate model to hold as follows.

(SR0) No arbitrage modelling. The short rate process r is such that for any $T \in [0,T^*]$ the discounted bond price process $(\tilde{P}(t,T) : t \in [0,T])$ is a martingale with respect to a measure $\mathbb{Q} \sim \mathbb{P}$.

(SR1) Bond price formula. There is an explicit formula for the bond prices $P(t,T)$ (in terms of the $\sigma(s,u)$ and by way of $P(t,T) = \mathrm{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^T r_u \mathrm{d}u\right)|\mathcal{F}_t\right]$).

These 2 properties are immediate from the construction of r_t (Corollary 5.9) resp. have been established in Corollary 5.10. Given a Lévy process L and a volatility structure σ we have a candidate for a short rate that satisfies (SR0) and (SR1) by the construction. This candidate should also satisfy the two following criteria of financial relevance.

(SR2) Positivity. Is the short rate positive, under which conditions?

(SR3) Mean reversion of r_t as expressed by existence of and convergence in probability to the random variable $r_\infty = \mathbb{P} - \lim_{t\to\infty} r_t$.

We do not have any criteria for the positivity of the short rate in a general case (for general volatility structure and driving process L). On the other hand we can construct criteria in terms of moment generating function telling us under which conditions the short rate process is mean reverting.

5.3.1. Distributional properties of the short rate As a preliminary for studying the mean convergence of the short rate we determine the moment generating function of r_t.

Proposition 5.11. *For any $z \in \mathbb{C}$ such that $z\sigma(s,t) \in \mathrm{Dom}(\psi_{L_1})$, and for any $s \in [0,t]$, we have*

$$\psi_{r_t}(z) = \exp(z\, m_t) \cdot \exp\left(\int_0^t \theta_{L_1}(z\sigma(s,t))\mathrm{d}s\right), \tag{5.14}$$

with the definition

$$m_t = f(0,t) + \theta_{L_1}(-\Sigma(0,t)).$$

Proof. The proposition here follows from (5.13) and the Key theorem 3.9, as

$$\psi_{r_t}(z) = \mathrm{E}\left[\exp\left(zm_t + \int_0^t z\sigma(s,t)\mathrm{d}L_s\right)\right] = \exp(z\, m_t)\,\mathrm{E}\left[\exp\left(\int_0^t z\sigma(s,t)\mathrm{d}L_s\right)\right].$$

□

5.3.2. Criteria for mean convergence of the short rate We study the existence and the distributional properties of

$$r_\infty \overset{\text{def}}{=} \mathbb{P} - \lim_{t\to\infty} r_t$$

at the level of moment generating functions.

Proposition 5.12. *Assume the existence of a function ψ_∞ on an open subset D_∞ of $\mathbb{C}$ containing a translate of the imaginary axis $\{\mathrm{i}y|y \in \mathbb{R}\}$ which is continuous near 0 and to which we have pointwise convergence of the moment generating functions of r_t with $t \uparrow \infty$. Then there is a random variable r_∞ such that*

$$\psi_\infty = \psi_{r_\infty} \qquad \text{on } D_\infty$$

and such that

$$r_\infty = \mathbb{P} - \lim_{t\to\infty} r_t.$$

The Proposition here follows from the Lévy continuity theorem; the precise result, quoted from (Bauer, 1996, Theorem 23.8), is as follows.

Theorem 5.13 (Lévy continuity theorem)**.** *Let (μ_n) be a sequence of finite (or bounded) Borel measures on $\mathbb{R}$. If the sequence of $(\hat{\mu}_n)$ of Fourier transforms converges pointwise to a complex-valued function φ on $\mathbb{R}$ that is continuous in 0, then φ is the Fourier transform of a (then uniquely determined) finite measure μ on $\mathbb{R}$ and $\mu_n \to \mu$ weakly for $n \to \infty$.*

We use this result in the following equivalent form.

Corollary 5.14. *Let (X_n) be a sequence of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with characteristic functions $\Psi_n = \Psi_{X_n}$. Assume that Ψ_n with $n \to \infty$ converges pointwise to a function Ψ on $\mathbb{R}$ which is continuous in a neigbourhood of 0. Then there exists a random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$ whose characteristic function is Ψ and to which the X_n converges in probability with $n \to \infty$.*

Proof of the Proposition 5.12. By application of the previous Corollary 5.14 with $X_n = r_n$, the assertion of the Proposition 5.12 is immediate. □

To render the result operational for determinig r_∞ we refer back to the exact form of the moment generating function for r_t, namely:

$$\psi_{r_t}(z) = \exp\left(z\, m_t + \int_0^t \theta_{L_1}(z\sigma(s,t))\mathrm{d}s\right)$$

with the definition

$$m_t = f(0,t) + \int_0^t \theta_{L_1}\left(-\Sigma(s,t)\right)\mathrm{d}s.$$

Sufficient for the existence of a pointwise limit of ψ_{r_t} for $t \to \infty$ is then the existence of such limits for each of the 2 sumands in their exponents. The principal form of such result is summarized as follows.

Corollary 5.15. *Assume the existence of the limit*

$$m_\infty \stackrel{\text{def}}{=} \lim_{t\to\infty} m_t$$

as well as the existence of a subset D_∞ of $\mathbb{C}$ on which we have existence of the limits

$$T_\infty(z) \stackrel{\text{def}}{=} \lim_{t\to\infty} \int_0^t \theta_{L_1}(z\sigma(s,t))\mathrm{d}s,$$

for any $z \in D_\infty$. If the resulting function given by

$$\psi_\infty(z) = \exp\left(m_\infty + T_\infty(z)\right), \qquad z \in D_\infty$$

is moreover continuous near $z = 0$, we have

$$\psi_\infty = \psi_{r_\infty},$$

where r_∞ is the limit of the r_t in probability:

$$r_\infty = \mathbb{P} - \lim_{t\to\infty} r_t.$$

Thus we have a machinery for the short rate and criteria that we have to apply for this short rate. The non-emptiness of such theory should be shown next, i.e. we ask about any example of a short rate produced by the machinery and satisfying (SR2) and (SR3). The postulated mean reversion of the short rate process indicates the possibility of modelling the short rate as an Ornstein–Uhlenbeck type process. These processes, studied in the Chapter 4, are mean reverting and thus taking r_t as an OU type process could be a possibility to achieve the mean reversion. The next section will in fact demonstrate the stylized facts for the OU type process case.

5.4 Short rate OU type process with drift

We will study the above properties of the short rate process and demonstrate them for a particular short rate model of generalized Vašíček type. In what follows we hence consider the Vašíček volatility structure

$$\text{(VOL)} \qquad \sigma(s,t) = \bar{\sigma}e^{-a(t-s)}, \qquad \text{for all } 0 \le s \le t \le T^*,$$

for real constants $a > 0$ and $\bar{\sigma} > 0$, unless expicitely stated otherwise.

Proposition 5.16. *Let σ follows the Vašíček volatility structure* (VOL)*. Then the no-arbitrage dynamics implied under (C0) – (C3) for the short rate process r is given by*

$$r_t = m(t) + \int_0^t \bar{\sigma}e^{-a(t-s)}\mathrm{d}L_s, \tag{5.15}$$

where

$$\begin{aligned} m(t) &= f(0,t) + \int_0^t \frac{\partial}{\partial t}\theta_{L_1}(-\Sigma(s,t))\mathrm{d}s = \\ &= f(0,t) + (-1)(\theta_{L_1}(0) - \theta_{L_1}(-\Sigma(0,t))) = \\ &= f(0,t) + \theta_{L_1}(-\Sigma(0,t)). \end{aligned}$$

Proof. The form of the short rate process follows from the Corollary 5.7 with specialization $\sigma(s,t) = \bar{\sigma}e^{-a(t-s)}$. Moreover, the integrated volatility structure

$$\Sigma(s,T) = \frac{\bar{\sigma}}{a}\left(1 - e^{-a(T-s)}\right)$$

satisfies the condition (S_ε) in the following form.

$$\frac{\partial}{\partial T}\Sigma(s,T) = \frac{\bar{\sigma}}{a}e^{-a(T-s)}(-a) = (-1)\frac{\partial}{\partial s}\Sigma(s,T),$$

hence $\varepsilon = (-1)$. Finally we have $\theta_{L_1}(0) = \log\psi_{L_1}(0) = 0$, as the characteristic function $\psi_{L_1}(0) = 1$. □

We may reexpress the short rate dynamics in the form of SDE as in the following result.

Corollary 5.17. *Assume moreover that $t \mapsto f(0,t)$ is continuously differentiable. Then the short rate process dynamics can be described via a SDE*

$$\mathrm{d}r_t = [\vartheta(t) - ar_t]\mathrm{d}t + \bar{\sigma}\mathrm{d}L_t, \tag{5.16}$$

where

$$\vartheta(t) = m'(t) + am(t),$$

or equivalently

$$r_t = r_0 + \int_0^t \vartheta(s)\mathrm{d}s + \bar{\sigma}L_t - a\int_0^t r_s\mathrm{d}s. \tag{5.17}$$

Proof. Differentiation with respect to time t of $X_t = \int_0^t \bar{\sigma}e^{-a(t-s)}\mathrm{d}L_s$ yields:

$$\mathrm{d}X_t = \left[-ae^{-at}\int_0^t \bar{\sigma}e^{as}\mathrm{d}L_s\right]\mathrm{d}t + \bar{\sigma}e^{-a(t-t)}\mathrm{d}L_t = a(m(t) - r_t)\mathrm{d}t + \bar{\sigma}\mathrm{d}L_t.$$

□

Corollary 5.18. *The short rate process is an Ornstein-Uhlenbeck type process with drift.*

5.4.1 Short rate GIG–OU model

In what follows we will assume that $(r_t : t \geq 0)$ is, in terms of Section 4.2, a GIG(λ,δ,γ)–OU type process and hence $(L_t : t \geq 0)$ is a BDLP of a GIG–OU type process. We will also consider a special case, where $(L_t : t \geq 0)$ is a BDLP of an IG(δ,γ)–OU type process. We have shown in the Lemma 4.17 that without loss of generatility we may assume that $\bar{\sigma} = 1$. In what follows we will thus assume that $\bar{\sigma} = 1$.

Distributional properties The log–moment generating functions were determined in the Section 5.3 and are given, respectively for the GIG and the IG case, by the following formulas.

$$\text{(GIG)} \qquad \theta_{L_1}(z) = \frac{\delta}{\gamma}\frac{az}{\sqrt{1-2z/\gamma^2}}\frac{K_{\lambda-1}(\delta\gamma\sqrt{1-2z/\gamma^2})}{K_\lambda(\delta\gamma\sqrt{1-2z/\bar{\gamma}^2})}, \qquad \mathrm{Re}(z) < \frac{\gamma^2}{2},$$

$$\text{(IG)} \qquad \theta_{L_1}(z) = \frac{\delta}{\gamma}\frac{az}{\sqrt{1-2z/\gamma^2}}, \qquad \mathrm{Re}(z) < \frac{\gamma^2}{2}.$$

It follows from the form of the domains of definition of θ_{L_1} in both the GIG and IG case, that L_1 satisfies the condition (C0).

As a preliminary for studying the properties (SR2) and (SR3) we will require a study of the distributional properties of r_t, in particular the determination of the law of r_t. Referring to the Proposition 5.11, the moment generating functions of r_t in the (GIG) and (IG) case are given as follows.

Proposition 5.19. *In the case (GIG), for arbitrary σ satysfying (C2) and (C4), we have for any $z \in \mathbb{C}$ such that* $\mathrm{Re}(z\sigma(s,t)) < \bar{\gamma}^2/2$

$$\psi_{r_t}(z) = \exp(z\, m(t)) \cdot \exp\left(\int_0^t \frac{\bar{\delta}}{\bar{\gamma}} \frac{az\sigma(s,t)}{\sqrt{1-2z\sigma(s,t)/\bar{\gamma}^2}} \frac{K_{\lambda-1}(\bar{\delta}\bar{\gamma}\sqrt{1-2z\sigma(s,t)/\bar{\gamma}^2})}{K_\lambda(\bar{\delta}\bar{\gamma}\sqrt{1-2z\sigma(s,t)/\bar{\gamma}^2})} \mathrm{d}s \right), \tag{5.18}$$

where

$$m_t = f(0,t) - \frac{\bar{\delta}}{\bar{\gamma}} \frac{a\Sigma(0,t)}{\sqrt{1+2\Sigma(0,t)/\bar{\gamma}^2}} \frac{K_{\lambda-1}\left(\bar{\delta}\bar{\gamma}\sqrt{1+2\Sigma(0,t)/\bar{\gamma}^2}\right)}{K_\lambda\left(\bar{\delta}\bar{\gamma}\sqrt{1+2\Sigma(0,t)/\bar{\gamma}^2}\right)}.$$

In the case (IG), this reduces to

$$\psi_{r_t}(z) = \exp(z\, m(t)) \cdot \exp\left(\int_0^t \frac{\bar{\delta}}{\bar{\gamma}} \frac{az\sigma(s,t)}{\sqrt{1-2z\sigma(s,t)/\bar{\gamma}^2}} \mathrm{d}s \right), \tag{5.19}$$

where

$$m(t) = f(0,t) - \frac{\bar{\delta}}{\bar{\gamma}} \frac{a\Sigma(0,t)}{\sqrt{1+2\Sigma(0,t)/\bar{\gamma}^2}}.$$

Proposition 5.20. *For $\sigma(s,t) = e^{-a(t-s)}$ we have in the case (GIG)*

$$\psi_{r_t}(z) = \exp\left(z\, m(t) + \theta_{GIG}(z) - \theta_{GIG}(ze^{-at})\right), \qquad \mathrm{Re}(z) < \frac{\gamma^2}{2}, \tag{5.20}$$

where

$$m_t \quad = \quad f(0,t) - \frac{\delta}{\gamma} \frac{(1-e^{-at})}{\sqrt{1+2a^{-1}(1-e^{-at})/\gamma^2}} \frac{K_{\lambda-1}(\delta\gamma\sqrt{1+2a^{-1}(1-e^{-at})/\gamma^2})}{K_\lambda(\delta\gamma\sqrt{1+2a^{-1}(1-e^{-at})/\gamma^2})}.$$

Corollary 5.21. *For $\sigma(s,t) = \bar{\sigma}e^{-a(t-s)}$ we have in the case (IG)*

$$\psi_{r_t}(z) = \exp\left(z\, m(t) + \delta\gamma\left(\sqrt{1-2ze^{-at}/\gamma^2} - \sqrt{1-2z/\gamma^2}\right)\right), \ \mathrm{Re}(z) < \frac{\gamma^2}{2}, \tag{5.21}$$

where

$$m(t) = f(0,t) - \frac{\delta}{\gamma} \frac{(1-e^{-at})}{\sqrt{1+2a^{-1}(1-e^{-at})/\gamma^2}}.$$

Proof of the Proposition 5.20. By application of the recurrence formula (A.6) for the K-Bessel functions we obtain

$$\frac{\partial}{\partial s} \log\left(\eta^\lambda(s,t) K_\lambda(\eta(s,t))\right) = \frac{1}{\eta^\lambda(s,t) K_\lambda(\eta(s,t))} \left(-\eta^\lambda(s,t) K_{\lambda-1}(\eta(s,t))\right) \frac{\partial}{\partial s}\eta(s,t),$$

setting

$$\eta(s,t) = \delta\gamma\sqrt{1 - 2za^{-1}e^{-a(t-s)}/\gamma^2}.$$

The derivative of $\eta(s,t)$ is obtained by partial differentiation with respect to the variable s and using the fact that

$$\frac{\partial}{\partial s} e^{-a(t-s)} = a e^{-a(t-s)}.$$

We thus obtain

$$\frac{\partial}{\partial s} \log\left(\eta^\lambda(s,t) K_\lambda(\eta(s,t))\right) = a\delta^2 \frac{z e^{-a(t-s)}}{\eta(s,t)} \frac{K_{\lambda-1}(\eta(s,t))}{K_\lambda(\eta(s,t))}. \tag{5.22}$$

Hence, using the previous result (5.22) together with (5.18) we obtain that

$$\begin{aligned}
\psi_{r_t}(z) &= \exp(z\, m(t)) \cdot \exp\left(\int_0^t \frac{\partial}{\partial s} \log\left(\eta^\lambda(s,t) K_\lambda(\eta(s,t))\right) \mathrm{d}s\right) \\
&= \exp(z\, m_t) \cdot \exp\left(\log\left(\eta^\lambda(t,t) K_\lambda(\eta(t,t))\right) - \frac{1}{a} \log\left(\eta^\lambda(0,t) K_\lambda(\eta(0,t))\right)\right) \\
&= \exp(z\, m_t) \cdot \exp\left(\log \psi_{GIG}(z) - \log \psi_{GIG}(z e^{-at})\right).
\end{aligned}$$

□

Although we are not able to determine the distribution of r_t, we can compute the moments. They may not in general determine the distribution of r_t, however they give us an idea about the behaviour of the short rate r_t. The first two moments are computed in the following Corollary.

Corollary 5.22. *The mean and variance of the short rate r_t are given by*

$$\begin{aligned}
\mathrm{E}\left[r_t\right] &= m_t + \mathrm{E}[GIG]\left(1 - e^{-at}\right) \\
\mathrm{var}\left[r_t\right] &= \mathrm{var}\left[GIG\right]\left(1 - e^{-2at}\right).
\end{aligned} \tag{5.23}$$

The moments of a random variable are computed using the following classic result, see i.e. (Lukacs, 1970, Theorem 2.3.1).

Lemma 5.23. *Let X be a random variable, with $X \in L_p$. Then*

$$\Psi_X^{(k)}(0) = \mathrm{i}^k \mathrm{E}[X^k] \qquad \textit{for all } k \le p.$$

Proof of Corollary 5.22. The characteristic function is given by $\Psi_{r_t}(u) = \psi_{r_t}(\mathrm{i}u)$ with $\psi_{r_t}(u)$ given by (5.14). By integration and some simple computation, using the Lemma 5.23, we obtain the result. □

Proposition 5.24. *The theoretical autocorrelation function of the short rate process r_t is given by*

$$\mathrm{corr}\left(r_t, r_{t+k}\right) = \frac{\mathrm{E}[GIG^2]\left(1 - e^{-2at}\right) - \mathrm{E}[GIG]^2\left(1 - e^{-at}\right)\left(1 - e^{-ak}\right)}{\mathrm{E}[GIG^2]\sqrt{\left(1 - e^{-at}\right)\left(1 - e^{-a(t+k)}\right)}}. \tag{5.24}$$

Proof. The correlation between r_t and r_{t+k} is given by

$$\mathrm{corr}(r_t, r_{t+k}) = \frac{\mathrm{E}[r_t r_{t+k}] - \mathrm{E}[r_t]\mathrm{E}[r_{t+k}]}{\sqrt{\mathrm{var}[r_t]\mathrm{var}[r_{t+k}]}}.$$

Since $r_t = m(t) + I_t$, where

$$I_t = \int_0^t e^{-a(t-s)} \mathrm{d}L_s,$$

it is enough to compute the moments of I_t and also $\mathrm{E}[I_t I_{t+k}]$. We have from (5.20)

$$\psi_{I_t}(z) = \exp\left(\theta_{GIG}(z) - \theta_{GIG}(ze^{-at})\right).$$

It follows from the previous Lemma 5.23 that

$$\begin{aligned}
\mathrm{E}[I_t] &= \mathrm{E}[GIG]\left(1 - e^{-at}\right)\\
\mathrm{E}[I_t^2] &= \mathrm{E}[GIG]^2\left(1 - e^{-at}\right)^2 + \mathrm{E}[GIG^2]\left(1 - e^{-2at}\right)\\
\mathrm{var}[I_t] &= \mathrm{E}[GIG^2]\left(1 - e^{-2at}\right)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{E}[I_t I_{t+k}] &= e^{-at}e^{-a(t+k)}\mathrm{E}\left[\int_0^t e^{as}\mathrm{d}L_s \int_0^{t+k} e^{as}\mathrm{d}L_s\right] =\\
&= e^{-at}e^{-a(t+k)}\mathrm{E}\left[\left(\int_0^t e^{as}\mathrm{d}L_s\right)^2\right] +\\
&\quad + e^{-at}e^{-a(t+k)}\mathrm{E}\left[\int_0^t e^{as}\mathrm{d}L_s \int_t^{t+k} e^{as}\mathrm{d}L_s\right].
\end{aligned}$$

It follows from the Theorem 3.10 about the independent incerements of the integrated Lévy process that the last term is equal to zero. We thus have

$$\mathrm{E}[I_t I_{t+k}] = e^{-ak}\mathrm{E}[I_t^2] = e^{-ak}\mathrm{E}[GIG]^2\left(1 - e^{-at}\right)^2 + \mathrm{E}[GIG^2]\left(1 - e^{-2at}\right).$$

The statement of the theorem then follows when we realize that

$$\mathrm{E}[r_t r_{t+k}] - \mathrm{E}[r_t]\mathrm{E}[r_{t+k}] = \mathrm{E}[I_t I_{t+k}] - \mathrm{E}[I_t]\mathrm{E}[I_{t+k}]$$

and $\mathrm{var}[r_t] = \mathrm{var}[I_t]$. □

Remark 5.25. In the GIG case and IG case, the first and second central moments are given by

$$\begin{aligned}
\mathrm{E}[GIG] &= \frac{\delta}{\gamma}\frac{K_{\lambda+1}(\delta\gamma)}{K_\lambda(\delta\gamma)}, &\qquad \mathrm{var}[GIG] &= \frac{\delta^2}{\gamma^2}\frac{K_{\lambda_2}(\delta\gamma)K_\lambda(\delta\gamma) + K_{\lambda+1}^2(\delta\gamma)}{K_\lambda^2(\delta\gamma)},\\
\mathrm{E}[IG] &= \frac{\delta}{\gamma}, &\qquad \mathrm{var}[IG] &= \frac{\delta}{\gamma^3}.
\end{aligned}$$

Positivity of the short rate The short rate r_t is given as a sum of a drift term m_t and a stochastic integral $\int_0^t \sigma(s,t)\mathrm{d}L_s$, with L_t being subordinator, i.e. a process with solely positive increments. It follows that the positivity of r_t depends only on the positivity of m_t. The precise result is as follows.

Proposition 5.26. *For any $t \geq 0$, the condition $m_t \geq 0$ is sufficient for $r_t \geq 0$.*

Proof. It follows from the Corollary 4.15 that the BDLP L is a subordinator, hence the stochastic integral

$$\int_0^t \bar{\sigma} e^{-a(t-s)} \mathrm{d}L_s$$

is always positive. Thus if $m_t \geq 0$, then also necessarily $r_t \geq 0$. □

Referring to the Corollary 5.18, the non-negativity of m_t leads to questions about weighted quotients of K-Bessel functions, since

$$m_t = f(0,t) - \delta^2 \frac{(1-e^{-at})}{\eta_t} \frac{K_{\lambda-1}(\eta_t)}{K_\lambda(\eta_t)},$$

where

$$\eta_t = \delta\gamma\sqrt{1 + 2a^{-1}\left(1-e^{-at}\right)/\gamma^2}.$$

For general λ we expect that the condition $m_t \geq 0$ has to be decided by numerical analysis. In the case (IG) with $\lambda = -1/2$, however, the situation is explicit as follows.

Proposition 5.27. *Suppose that the short process r_t is an IG–OU type process given by*

$$r_t = m(t) + \int_0^t e^{-a(t-s)} \mathrm{d}L_s,$$

with

$$m(t) = f(0,t) - \frac{\delta}{\gamma} \frac{\left(1-e^{-at}\right)}{\sqrt{1 + 2a^{-1}\left(1-e^{-at}\right)/\gamma^2}}.$$

Assume that $f(0,t) > 0$ for all $t \geq 0$. Then, for any $a > 0$ and any $t \geq 0$, the following statements are equivalent.

(i) $m_t \geq 0$.

(ii) $\Sigma_t = \frac{1}{a}(1-e^{-at})$ *belongs to a solution interval for the quadratic equation*

$$a^2\delta^2 x^2 - 2f^2(0,t)x - \gamma^2 f^2(0,t) \leq 0. \tag{5.25}$$

(iii)

$$1 < \frac{a}{(1-e^{-at})}\Sigma_t^+, \tag{5.26}$$

where

$$\Sigma_t^+ = \frac{1}{a^2\delta^2}\left(f^2(0,t) + f(0,t)\sqrt{f^2(0,t) + a^2\delta^2\gamma^2}\right)$$

is the positive root of the quadratic polynomial (5.25)*.*

Proof. By solving the equation $m(t) \geq 0$ for Σ_t we obtain the quadratic equation as in (ii). By putting the LHS equal to zero we obtain the zero points

$$\Sigma_t^{\pm} = \frac{1}{a^2\delta^2}\left(f^2(0,t) \pm f(0,t)\sqrt{f^2(0,t) + a^2\delta^2\gamma^2}\right).$$

Since $\Sigma_t^- < 0 < \Sigma_t$ we obtain a condition for positivity of the drift $m(t)$ in the form (5.26). □

Corollary 5.28. *With the notation of the previous Proposition 5.27 we then have*

$$\text{(POS)} \qquad r_t > 0 \qquad \textit{for any } t \in [t_0, T^*] \subseteq [0, T^*]$$

if the following two conditions are satisfied:

(i) $f(0,t) > 0$ *for all* $t \in [t_0, T^*]$,

(ii)

$$1 < \frac{a}{(1 - e^{-at})} \min_{t \in [t_0, T^*]} \Sigma_t^+.$$

Mean convergence of the short rate We apply the criteria of the Corollary 5.15 to the case when r_t is GIG–OU type process with Vašíček volatility structure (VOL).

Theorem 5.29. *Assume the existence of the limit*

$$f(0,\infty) = \lim_{t\to\infty} f(0,t).$$

Then we have on $D_\infty = \{z \in \mathbb{C} | \mathrm{Re}(z) \leq \gamma^2/2\}$ *pointwise convergence of the moment generating functions* ψ_{r_t} *with* $t \to \infty$ *to the continuous function* ψ_∞ *on* D_∞ *given by*

$$\psi_\infty(z) = \exp\left(z\, m_\infty + T_\infty(z)\right),$$

where

$$m_\infty = \lim_{t\to\infty} m_t = f(0,\infty) - \frac{\delta}{\gamma}\frac{1}{\sqrt{1 + 2/(\gamma^2 a)}}\frac{K_\lambda(\delta\gamma\sqrt{1 + 2/(\gamma^2 a)})}{K_\lambda(\delta\gamma)}$$

and

$$T_\infty(z) = \lim_{t\to\infty} \theta_{GIG}(z) - \theta_{GIG}(ze^{-at}) = \theta_{GIG}(z)$$

In the IG case the precise result is as follows.

Corollary 5.30. *In the situation of Theorem 5.29, in the IG case we have*

$$\psi_\infty(z) = \exp\left(z\, m_\infty + T_\infty(z)\right),$$

where

$$m_\infty = \lim_{t\to\infty} m_t = f(0,\infty) - \frac{\delta}{\gamma}\frac{1}{\sqrt{1 + 2/(\gamma^2 a)}},$$

and

$$T_\infty(z) = \lim_{t\to\infty} \delta\gamma\left(\sqrt{1 - 2ze^{-at}/\gamma^2} - \sqrt{1 - 2z/\gamma^2}\right) = \delta\gamma\left(1 - \sqrt{1 - 2z/\gamma^2}\right).$$

Corollary 5.31. *Under the conditions of Corollary 5.29, we have the existence of the limit in probability of the short rate r_t with $t \to \infty$,*

$$r_\infty = \mathbb{P} - \lim_{t \to \infty} r_t,$$

which has GIG$(\lambda, \delta, \gamma)$ distribution shifted by m_∞.

5.4.2 Short rate OU–NIG model

In the previous section we studied a model of the short rate, where the driving process was a subordinator. This assumption leads to a model of a short rate that can only jump upwards. However, it is often the case that the short rate process has negative jumps also. It makes sense to study a model, where this property is included. One of the possibilities is to consider a general Lévy process as the driving process instead of just a subordinator. Eberlein and Raible (1999) consider hyperbolic processes and Eberlein and Kluge (2007) consider more generally generalized hyperbolic processes. In what follows we will focus our attention on NIG processes and assume that the short rate is an OU–NIG type process, i.e., the BDLP is a NIG$(\alpha, \beta, \mu, \delta)$ process.

Distributional properties The log–moment generating function of L_1 is given by

$$\text{(NIG)} \qquad \theta_{L_1}(z) = \delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + z)^2}\right) + \mu z, \qquad \mathrm{Re}(|z + \beta|) < \alpha.$$

First we determine the law of r_t. Referring to the Proposition 5.11, the moment generating functions of r_t in the (NIG) case is given as follows.

Proposition 5.32. *In the case (NIG), for arbitrary σ satysfying (C2) and (C4), we have for any $z \in \mathbb{C}$ such that* $\mathrm{Re}(|z\sigma(s,t) + \beta|) < \alpha$

$$\psi_{r_t}(z) = \exp\left(z\, m(t) + \int_0^t \left[\delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + z\sigma(s,t))}\right) + z\,\mu\,\sigma(s,t)\right] \mathrm{d}s\right), \tag{5.27}$$

where

$$m(t) = f(0,t) - \delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta - \Sigma(0,t))^2}\right) - \mu\Sigma(0,t).$$

Let us assume for simplicity that $\beta = 0$. Then the previous result takes with the Vašíček volatility structure the following specific form.

Proposition 5.33. *For $\sigma(s,t) = \bar\sigma e^{-a(t-s)}$ and $\beta = 0$ we have in the case (NIG)*

$$\psi_{r_t}(z) = \exp\left(z\, m_t + \mu\, z\frac{\bar\sigma}{a}\left(1 - e^{-at}\right) - \frac{\delta}{a}\sqrt{\alpha^2 - z^2\bar\sigma^2} + \right.$$
$$\left. + \frac{\delta}{a}\sqrt{\alpha^2 - z^2\bar\sigma^2 e^{-2at}} + \frac{\delta}{\gamma}\alpha \log\frac{\sqrt{\alpha^2 - z^2\bar\sigma} + \alpha}{\sqrt{\alpha^2 - z^2\bar\sigma^2 e^{-2at}} + \alpha}\right), \qquad \mathrm{Re}(\bar\sigma|z|) < \alpha,$$

where

$$m_t \;=\; f(0,t) - \delta\left(\alpha - \sqrt{\alpha^2 - \bar\sigma^2(1 - e^{-at})^2/a}\right) - \mu\frac{\bar\sigma}{a}\left(1 - e^{-at}\right).$$

Positivity of the short rate As the driving process has NIG distribution, the short rate process can be negative with positive probability. However, computation of this probability is quite involved, as we do not know the density of the short rate process and has to be solved numerically.

Mean convergence We apply the criteria of the Corollary 5.15 to the case when r_t is OU–NIG type process with Vašíček volatility structure (VOL).

Theorem 5.34. *Assume the existence of the limit*

$$f(0,\infty) = \lim_{t\to\infty} f(0,t).$$

Then we have on $D_\infty = \{z \in \mathbb{C} | \operatorname{Re}(|z|) < \alpha\}$ pointwise convergence of the moment generating functions ψ_{r_t} with $t \to \infty$ to the continuous function ψ_∞ on D_∞ given by

$$\psi_\infty(z) = \exp\left(z\, m_\infty + T_\infty(z)\right),$$

where

$$m_\infty = \lim_{t\to\infty} m_t = f(0,\infty) + \delta\alpha - \delta\sqrt{\alpha^2 - \frac{\bar{\sigma}^2}{a^2}} - \mu\frac{\bar{\sigma}}{a}$$

and

$$\begin{aligned} T_\infty(z) = \lim_{t\to\infty} \mu\, z\frac{\bar{\sigma}}{a}\left(1 - e^{-at}\right) - \frac{\delta}{a}\sqrt{\alpha^2 - z^2\bar{\sigma}^2} + \frac{\delta}{a}\sqrt{\alpha^2 - z^2\bar{\sigma}^2 e^{-2at}} + \\ + \frac{\delta}{\gamma}\alpha \log\frac{\sqrt{\alpha^2 - z^2\bar{\sigma}} + \alpha}{\sqrt{\alpha^2 - z^2\bar{\sigma}^2 e^{-2at}} + \alpha} = \\ = z\,\mu\frac{\bar{\sigma}}{a} - \frac{\delta}{a}\sqrt{\alpha^2 - z^2\bar{\sigma}^2} + \frac{\delta}{a}\alpha\left(1 + \log\frac{\sqrt{\alpha^2 - z^2\bar{\sigma}^2} + \alpha}{2\alpha}\right). \end{aligned}$$

Chapter 6

Model calibration

In this chapter we will describe a method of estimations the parameters of the term structure model described in the previous Chapter 5. We will focus on the OU–NIG model of the short rate, following the work of Raible (2000) and Eberlein and Kluge (2007).

There are several ways how to estimate the parameters of the model. Eberlein and Kluge (2007) describe two ways of so called calibration of the model to the market data.

Calibration under the real-world measure Here we use the market data of zero coupon bond prices or yield curves to obtain a random sample from our distribution. By statistical methods such as maximum likelihood or generalized method of moments we estimate the parameters of the distribution.

Calibration under the risk-neutral measure In this method we try to fit the model implied prices of interest rate derivatives such as caps and swaptions to the prices observed on the market. Eberlein and Kluge (2006) derived exact pricing formulae for caps and swaption in the Lévy driven term structure model described in the previous chapter. The parameters estimates are obtained by minimizing the distance between these model prices and the market prices given by implied volatilities, over all possible values of the unknown parameters.

We will use the first method, i.e. the calibration under the real-world measure (that coincides in our model with the risk-neutral measure in our model).

6.1 Parameters estimation in the OU–NIG model of short rate

Consider the OU–NIG short rate model from the Subsection 5.4.2, where the short rate process is given by

$$r_t = m(t) + \int_0^t \bar{\sigma} e^{-a(t-s)} \mathrm{d}L_s, \tag{6.1}$$

where

$$m(t) = f(0,t) + \theta_{L_1}(-\Sigma(0,t)),$$

and the driving Lévy process is NIG process. This model allows negative jumps of the BDLP, implying a possibility of the short rate to be negative.

The parameters that are unknown and have to be estimated are a, σ and parameters of the NIG distribution α, β, μ and δ. We use the method described in Eberlein and Kluge (2007) and Raible (2000) to obtain a sample of realizations of the background Lévy process L_1 from the zero coupon bond prices or the yield curves. Then we will use the maximum likelihood to estimate the parameters of L_1 (i.e. the parameters of the NIG distribution).

6.1.1 Random sample extraction

For every trading day between April 3rd, 2006 and November 20, 2009 (in total 929 trading days) we have a list of yield curves from the Euro area with maturities from 1 day, 1 week, 1 month, 2 months, 3 months, 6 months, 9 months, 1 year up to 10 years and 15 years. From this data we compute the prices of zero coupon bonds on every day by use of the following formula

$$P(t,T) = \exp\{-(T-t)y(t,T)\},$$

where $y(t,T)$ is the value of the yield curve for maturity T at time t. The yield curve from November 20, 2009 is showed in Figure 6.1.

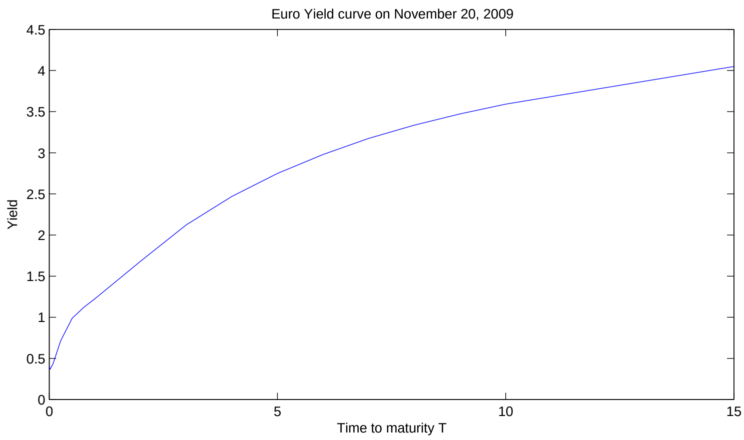

Figure 6.1: Euro yield curve on November 20, 2009.

We define the daily log return from the zero coupon bond with maturity T over the time period $[t, t+1)$ (t is in days) as the logarithm of ratio between the bond price and its forward price on the day before, that is equal to

$$LR(t,T) = \log \frac{P(t+1, t+T)}{P(t, t+1, t+T)} = \log P(t+1, t+T) - \log P(t, t+T) + \log P(t, t+1).$$

We determine the daily log returns from the data. For $k \in \{0, 1, \dots, N\}$ with $N = 929$ and $T \in \{1D, 1M, \dots, 15Y\}$ years to maturity we compute $LR(k, k+T)$. However, from the data available we are not able to extract $P(k, k+1)$ and $P(k+1, k+T)$, thus we use the interpolation method described in Raible (2000) and we interpolate the logarithm of the bond prices with a cubic spline. The interpolated logarithm of the bond prices from November 20, 2009 is showed in Figure 6.2.

We can rewrite the logaritmic returns using (5.6) to obtain

$$LR(t,T) = -\int_t^{t+1} A(s, t+T) - A(s, t+1)\mathrm{d}s - \int_t^{t+1} \Sigma(s, t+T) - \Sigma(s, t+1)\mathrm{d}L_s.$$

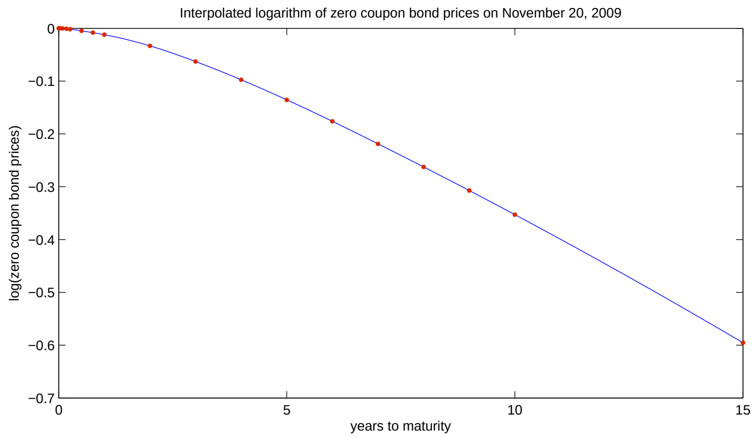

Figure 6.2: Interpolated logarithm of zero coupon bond prices on November 20, 2009.

It follows from the stationarity of the volatility structure $\sigma(s,t) = \sigma(0, t-s)$ that we may write

$$LR(t,T) = d(T) + c(T)Y_{t+1},$$

where $d(T)$ and

$$c(T) = -\frac{\sigma}{a}\left(1 - e^{-a(T-1)}\right)$$

are deterministic functions and $Y_{t+1} = L_{t+1} - L_t \sim L_1$. Without loss of generality we may assume that $\sigma = 1$. The parameter a has to be estimated by other methods, here we will assume that $a = 0.01$. Next assume that $\mathrm{E}[L_1] = 0$. Then we have

$$LR(t,T) - \mathrm{E}[LR(t,T)] = c(T)Y_{t+1}. \tag{6.2}$$

$Y_1, \ldots, Y_N$ are independent and equal to L_1 in distribution and the sample $y_1, \ldots, y_N$ corresponding to $Y_1, \ldots, Y_N$ can be obtained in the following way. We realize that (6.2) implies that y_{k+1} can be found by solving linear regression problem $\mathbf{y} = \beta \mathbf{X}$ with

$$\begin{aligned}
\mathbf{y} &= \left(LR(k, k+1D) - \overline{LR}_{1D}, \ldots, LR(k, k+15Y) - \overline{LR}_{15Y}\right)', \\
\mathbf{X} &= (c(1D), \ldots, c(15Y))', \\
\beta &= y_{k+1}
\end{aligned}$$

for every $k = 1, \ldots, N$ and where

$$\overline{LR}_T = \frac{1}{N}\sum_{k=1}^{N} LR(k, k+T).$$

In this way we obtain the sample $y_1, \ldots, y_N$ which can now be used to estimate the parameters of the driving process L_1. The empirical density function of the random variable L_1 against the Gaussian density function is in Figure 6.3. The empirical density was obtained using the kernel estimator in MATLAB. It is clear that the Gaussian distribution does not fit the data well.

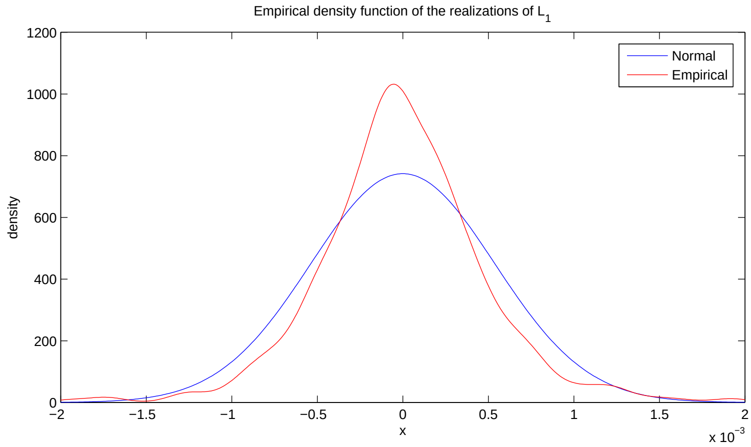

Figure 6.3: Empirical density of the random variable L_1.

6.1.2 Maximum likelihood estimation

We will use the maximum likelihood to obtain the parameters of the driving process L_1 that is assumed to be NIG$(\alpha, \beta, \mu, \delta)$ distributed. The density function of L_1 is given by

$$f(x) = \frac{\alpha\delta}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)\right) \frac{K_1(\alpha\sqrt{\delta^2 + (x-\mu)^2})}{\sqrt{\delta^2 + (x-\mu)^2}}, \qquad x \in \mathbb{R}.$$

The log likelihood function is then given by

$$\ell(\alpha, \beta, \mu, \delta | y_1, \ldots, y_N) = N \log \frac{\alpha\delta}{\pi} + N\delta\sqrt{\alpha^2 - \beta^2} + \sum_{i=1}^{N} \left\{ \beta(y_i - \mu) + \log K_1\left(\alpha\sqrt{\delta^2 + (y_i - \mu)^2}\right) - \log\sqrt{\delta^2 - (y_i - \mu)^2} \right\}.$$

By the maximization of this likelihood function we obtained the estimated paramaters that are given in the Table 6.1. Figure 6.4 shows the estimated density function and the empirical density function together with the Gaussian distribution.

α	β	μ	δ
1571.81	126.16	4.25e-04	-3.42e-05

Table 6.1: Estimated parameters of the NIG distribution

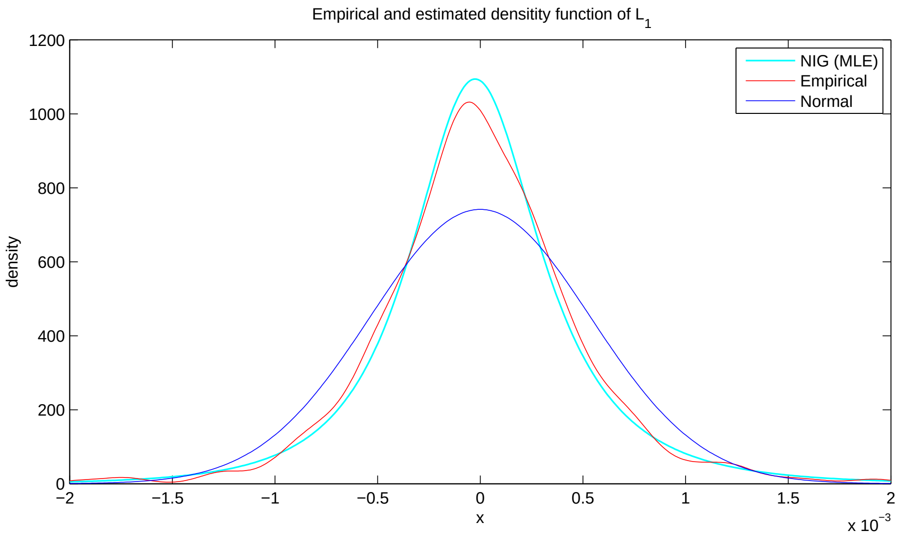

Figure 6.4: Estimated and empirical density of the random variable L_1.

Chapter 7

Numerical methods

The Lévy model presented in the Chapter 5 may be used in various ways, from pricing of interest rate derivatives to risk management. There are several analytical pricing methods for caps and swaptions, see Eberlein and Kluge (2006), however for other application, such as pricing of exotic options, scenario simulation for risk management, etc., we need numerical methods.

Assuming the risk-neutral valuation approach to hold, the principal problem to be adressed is computing a price at time t for an interest rate derivative with value $C_T = C\left((r_s, s \in [0,T])\right)$ at time T by way of computing the conditional expectation

$$C_t = \mathrm{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s \mathrm{d}s\right) C_T|\mathcal{F}_t\right], \tag{7.1}$$

where $\mathbb{Q}$ is the equivalent martingale measure.

There are several ways how to compute the C_t in practice.

(1) **PDE technique** Using the Markov property of r_t and the Feynman-Kac theorem, the price at time t of the contingent claim C_t turns out to be a solution of a partial differential equation.

(2) **Change of numeraire** By choosing an appropriate numeraire, the computation of the conditional expectation gets easier.

(3) **Monte Carlo simulation** We approximate the expectation $\mathrm{E}_{\mathbb{Q}}[\exp\left(-\int_0^T r_s \mathrm{d}s\right) C_T]$ by the arithmetic mean of a large number of realizations of $\exp\left(-\int_0^T r_s \mathrm{d}s\right) C_T$. By the law of large numbers, this mean converges to the desired expectation.

(4) **Lattice approximation** We assume that we have a sequence r^n of stochastic processes weakly converging to r. Then, under some technical conditions also

$$\mathrm{E}_{\mathbb{Q}}\left[\exp\left(-\int_0^T r_s^n \mathrm{d}s\right) C(r^n)\right] \to \mathrm{E}_{\mathbb{Q}}\left[\exp\left(-\int_0^T r_s \mathrm{d}s\right) C(r)\right].$$

If we have some lattice approximation r^n of the process r then the expectation on the left-hand side can be computed by the backward induction, based on the rule of iterated conditional expectations and the Markov property of r^n.

Here we will treat the third and the fourth method.

7.1 Simulations

This section reviews the method of exact simulation of the IG–OU type processes based on the Proposition 4.23 and the work Zhang and Zhang (2008). The algorithms for simulation of IG–OU type processes as well as the IG–OU type short rate process are described.

7.1.1 Exact simulation of the IG–OU type process

Let $a > 0$, $\sigma = 1$ and consider an IG–OU type process as defined by (4.3), i.e.

$$X_t = e^{-at} X_0 + \int_0^t e^{-a(t-s)} \mathrm{d}L_s.$$

From the Proposition 4.23 we know, that the stochastic integral

$$I_t = \int_0^t e^{-a(t-s)} \mathrm{d}L_s$$

can be represented as a sum of an inverse Gaussian random variable and a compound Poisson random variable, $I_t = I_t^{(1)} + I_t^{(2)}$, where $I_t^{(1)}$ is $\mathrm{IG}\Big(\delta\left(1 - e^{-\frac{1}{2}at}\right), \gamma\Big)$ random variable and $I_t^{(2)}$ is a compound Poisson process with intensity $\delta\left(1 - e^{-\frac{1}{2}at}\right)$ and with jumps having common density function

$$f_t(x) = \begin{cases} \frac{\gamma^{-1}}{\sqrt{2\pi}} x^{-3/2} \left(e^{\frac{1}{2}at} - 1\right)^{-1} \left(e^{-\frac{1}{2}\gamma^2 x} - e^{-\frac{1}{2}\gamma^2 x e^{at}}\right), & x > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{7.2}$$

We know from the Remark 4.7 that an OU type process is a temporally homogenous Markov process. Hence, for a time interval of length Δ and a starting point $X_t = x$, we can write

$$X_{t+\Delta} = e^{-a\Delta} x + \int_0^\Delta e^{-a(\Delta-s)} \mathrm{d}L_s = e^{-a\Delta} x + I_\Delta = e^{-a\Delta} x + I_\Delta^{(1)} + \sum_{i=1}^{N_\Delta} W_i^\Delta. \tag{7.3}$$

To obtain a simulated IG–OU type process $(X(j\Delta), j = 0, 1, \dots)$ we thus need to simulate IG random variables (Algorithm 7.1) and random variables from distribution having density function $f_\Delta(x)$ (Algorithm 7.2).

Algorithm 7.1. Generation of the $\mathrm{IG}(\delta, \gamma)$ random variable can be done in the following way (Michael et al. (1976))

1. Generate a χ_1^2 random variable Y.
2. Set $X_1 = \frac{\delta}{\gamma} + \frac{Y}{2\gamma^2} - \frac{1}{2\gamma^2}\sqrt{4\delta\gamma Y + Y^2}$.
3. Generate uniform $U(0,1)$ random variable U. If $U \le \frac{\delta}{\delta + \gamma X_1}$ set $X = X_1$. Otherwise set $X = \frac{\delta^2}{\gamma^2 X_1}$.

Then $X \sim IG(\delta, \gamma)$.

Algorithm 7.2. Generation of the random variables from the density $f_\Delta(x)$ (Zhang and Zhang (2008))

1. Generate a $\Gamma\left(\frac{1}{2}, \frac{1}{2}\gamma^2\right)$ random variable Y.

2. Generate a uniform $U(0,1)$ random variable U.

3. If $U \leq \frac{f_\Delta(Y)}{\frac{1}{2}\left(1+e^{\frac{1}{2}a\Delta}\right)g(Y)}$, set $X = Y$, where $g(y) = \frac{\left(\frac{1}{2}\gamma^2\right)^{1/2}}{\Gamma\left(\frac{1}{2}\right)} y^{-\frac{1}{2}} e^{-\frac{1}{2}\gamma^2 y}$. Otherwise return to 1.

Then X is from density $f_\Delta(x)$.

The initial value X_0 is taken from the invariant density IG(δ, γ). Once we have generated a random variable $X(t)$ it follows from (7.3) that the next value $X(t + \Delta)$ can be simulated in the following way.

Algorithm 7.3. Generation of random variable $X(t+\Delta)$ given the value $X(t)$ (Zhang and Zhang (2008))

1. Generate a IG$\left(\delta\left(1 - e^{-\frac{1}{2}a\Delta}\right), \gamma\right)$ random variable $I_t^{(1)}$ using the Algorithm 7.1.

2. Generate a Poisson Po$\left(\delta\gamma\left(1 - e^{-\frac{1}{2}a\Delta}\right)\right)$ random variable N_Δ.

3. Generate $W_1^\Delta, \ldots, W_{N_\Delta}^\Delta$ independent random variables from density $f_\Delta(x)$ using the Algorithm 7.2.

4. Set
$$X(t + \Delta) = e^{-a\Delta} X(t) + I_\Delta^{(1)} + \sum_{i=1}^{N_\Delta} W_i^\Delta.$$

Then $X(t + \Delta)$ is the simulated value of IG–OU type process at time $t + \Delta$.

Example 7.4. Here we give an example of such simulation that was implemented in MATLAB. Examples of simulated IG–OU type process for different parameters are given in Figure 7.1. From the figure we can evidence the fact that an IG–OU type process moves up entirely by jumps and then tails off exponentially.

The process (X_t) is time-homogenous and for any $t > 0$ and $s > 0$ the correlation between X_t and X_{t+s} is given by
$$r(s) = \mathrm{corr}(X_t, X_{t+s}) = e^{-as}.$$
The simulated data gives evidence to this fact as can be seen from Figure 7.2. This figure represents the empirical autocorrelation function (the correlogram on the picture) versus the theoretical autocorrelation function $r(s)$ (the solid line). The difference is insignificant.

We can also exploit to what extent the simulated data really comes from the IG distribution. From the simulated data we obtain the empirical density function and compare it with the original density function of the IG distribution, that is given by

$$p_{IG}(x) = \frac{1}{\sqrt{2\pi}} \delta e^{\delta\gamma} x^{-3/2} \exp\left\{-\frac{1}{2}(\delta^2 x^{-1} + \gamma^2 x)\right\}, \qquad x \in \mathbb{R}^+.$$

The results, that are in Figure 7.3, show that the IG distribution is simulated quite well. The empirical function was obtained by the kernel density estimation function.

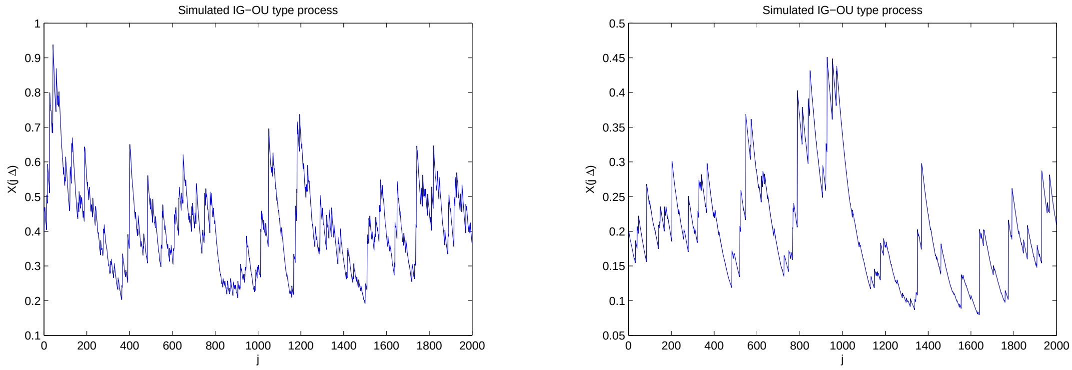

Figure 7.1: Simulated IG-OU type process $X(j\Delta)$ against j. Left: $a = 0.02, \delta = 2, \gamma = 5$ and $\Delta = 1$. Right: $a = 0.01, \delta = 1, \gamma = 5$ and $\Delta = 1$.

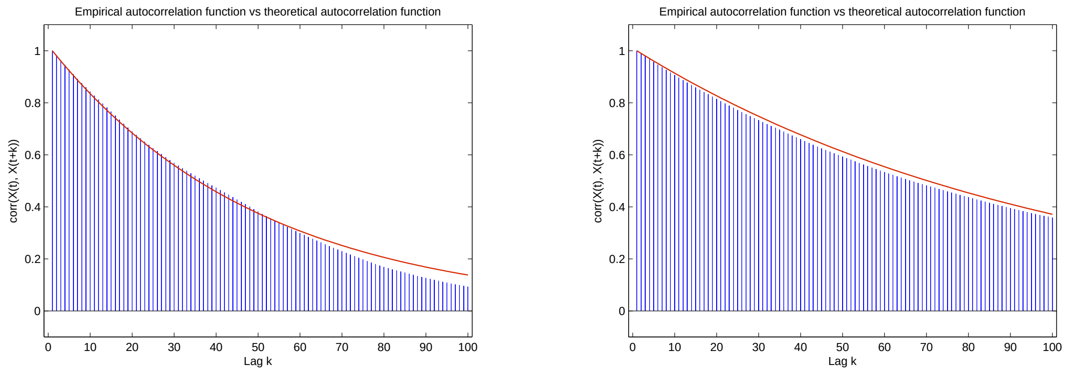

Figure 7.2: Empirical autocorrelation function of the simulated data $X(j\Delta)$, $j = 0, \ldots, 20000$, vs. the theoretical autocorrelation function. Left: $a = 0.02, \delta = 2, \gamma = 5$ and $\Delta = 1$. Right: $a = 0.01, \delta = 1, \gamma = 5$ and $\Delta = 1$.

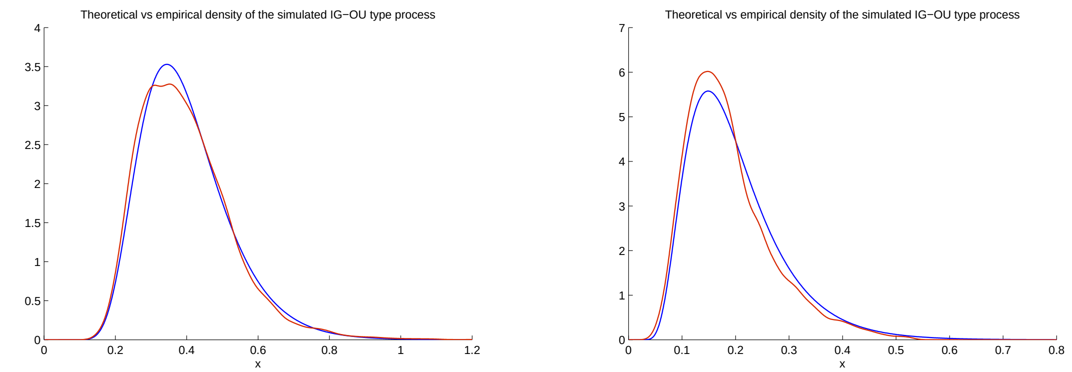

Figure 7.3: Empirical (red) and theoretical (blue) density function of the simulated IG–OU type process. Left: $a = 0.02, \delta = 2, \gamma = 5$ and $\Delta = 1$. Right: $a = 0.01, \delta = 1, \gamma = 5$ and $\Delta = 1$.

7.1.2 Simulation of the short rate IG–OU type process

In this Subsection we describe algorithm to simulate the short rate process from the Subsection 5.4.1, i.e. a short rate that is a IG(δ, γ)–OU type process with drift. Since we have algorithm to simulate the IG–OU type process, the task to simulate the short rate process r_t becomes quite easy. The short rate process r_t satisfies the SDE

$$\begin{aligned} r_t &= e^{-a(t-s)} r_s + \int_0^t e^{-a(t-u)} \vartheta(u) \mathrm{d}u + \int_s^t \bar{\sigma} e^{-a(t-u)} \mathrm{d}L_u = \\ &= e^{-a(t-s)} r_s + a m(t) - a e^{-a(t-s)} m(s) + \int_s^t \bar{\sigma} e^{-a(t-u)} \mathrm{d}L_u. \end{aligned} \tag{7.4}$$

The initial value $r_0 = m(0)$ is a deterministic value. Once we have generated a random variable r_t it follows from (7.4) that the next value $r_{t+\Delta}$ can be simulated in the following way.

Algorithm 7.5. Generation of random variable $r_{t+\Delta}$ given the value r_t

1. Let $\bar{\delta} = \delta \sigma^{-1/2}$ and $\bar{\gamma} = \gamma \sigma^{1/2}$.
2. Generate a IG$\left(\bar{\delta}\left(1 - e^{-\frac{1}{2}a\Delta}\right), \bar{\gamma}\right)$ random variable $I_t^{(1)}$ using the Algorithm 7.1.
3. Generate a Poisson Po$\left(\bar{\delta}\bar{\gamma}\left(1 - e^{-\frac{1}{2}a\Delta}\right)\right)$ random variable N_Δ.
4. Generate $W_1^\Delta, \ldots, W_{N_\Delta}^\Delta$ independent random variables from density $f_\Delta(x)$ using the Algorithm 7.2.
5. Set
$$r_{t+\Delta} = e^{-a\Delta} r_t + a m(t+\Delta) - a e^{-a\Delta} m(t) + I_\Delta^{(1)} + \sum_{i=1}^{N_\Delta} W_i^\Delta.$$

Then $r_{t+\Delta}$ is the simulated value of IG–OU type process with drift at time $t + \Delta$.

Example 7.6. A result of one such simulation implemented in MATLAB is given in Figure 7.4. One can see that the positivity of the short rate is satisfied, as well as the mean convergence (return to the mean of the process).

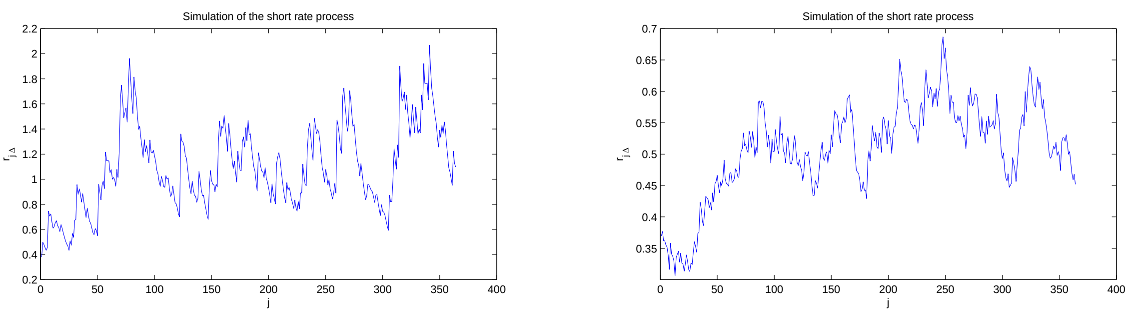

Figure 7.4: Simulated IG–OU short rate process. Left: $a = 0.1, \sigma = 0.1, \delta = 1, \gamma = 10$ and $\Delta = 1$. Right: $a = 0.1, \sigma = 0.5, \delta = 4, \gamma = 20$ and $\Delta = 1$.

For comparison we simulated also a short rate proces in the diffusion context, where the Lévy process is simply a Brownian motion. This model corresponds to the classical Vašíček model. A result of such simulation is in Figure 7.5.

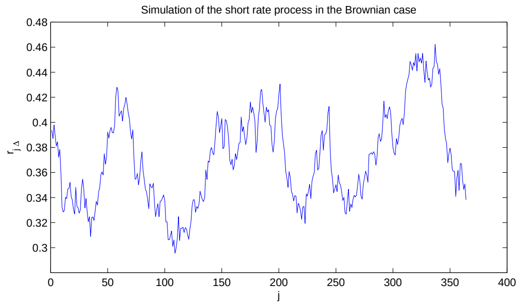

Figure 7.5: Simulated short rate process in the Brownian case. $a = 0.01, \sigma = 0.1$, and $\Delta = 1$.

7.1.3 Simulation of the short rate OU–NIG type process

In this Subsection we describe algorithm to simulate the short rate process from the Subsection 5.4.2., i.e. a short rate that is an OU–NIG$(\alpha, \delta, \mu, \delta)$ type process with drift. We remind (see Subsection 2.4.7) that the normal inverse Gaussian process can be obtained by time-changing a standard Brownian motion with drift by an IG process. Hence a simulation of such process involves simulation of standard normal random variables and IG distributed random variables.

Algorithm 7.7. Generation of NIG$(\alpha, \beta, \mu, \delta)$ random variable can be done in the following way (Rydberg (1997))

1. Generate a IG$(\delta, \sqrt{\alpha^2 - \beta^2})$ random variable Z.
2. Generate a standard normal N$(0, 1)$ random variable Y.
3. Set $X = \mu + \beta Z + \sqrt{Z} Y$.

Then $X \sim$ NIG$(\alpha, \beta, \mu, \delta)$

We use this algorithm for simulation of OU–NIG short rate process. We remind that the short rate process satisfies the SDE,

$$r_t = e^{-a(t-s)} r_s + am(t) - ae^{-a(t-s)} m(s) + \int_s^t \sigma e^{-a(t-u)} \mathrm{d}L_u.$$

For $(t - s) \to 0$ we may approximate the stochastic integral by the Riemann sum.

Algorithm 7.8. Generation of random variable $r_{t+\Delta}$ given the value r_t.

1. Generate a NIG$(\alpha, \beta, \Delta\mu, \Delta\delta)$ random variable L_Δ.
2. Set
$$r_{t+\Delta} = e^{-a\Delta} r_t + am(t + \Delta) - ae^{-a\Delta} m(t) + \sigma e^{-a\Delta} L_\Delta.$$

Then $r_{t+\Delta}$ is the simulated value of OU–NIG type process with drift at time $t + \Delta$.

Example 7.9. A result of such simulation with the estimated parameters from the Section 6.1 is given in Figure 7.6. One can directly see the disadvantage of this model, the negative short rates.

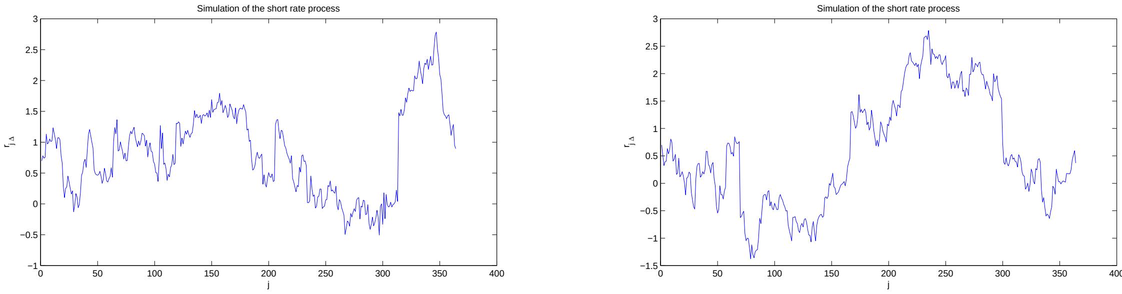

Figure 7.6: Simulated OU–NIG short rate process with parameters from Table 6.1, with $a = 0.01$, $\sigma = 1$ and $\Delta = 1$.

7.2 Lattice method

This section develops an extension of the Schmidt (1997) approach of the valuation of contingent claim using the lattice approximations to the Lévy driven interest rates. This method consist of construction of multinomial trees of the considered process, in our case the short rate process, and use them for pricing path-dependent interest rate derivatives such as american and bermudian options. In the classical diffusion approach this method is very popular, since approximation of the Gaussian random variables is quick and simple.

In Chapter 5 we developed a term structure model driven by a Lévy process. In this Section now we more precisely develop a discrete approximation of the short rate process in the IG–OU case. As a preliminary to this, we develop approximations to the compound Poisson and IG processes, that are cornerstones for the short rate process approximation.

7.2.1 Approximation of Lévy processes

Lattice approximation of Poisson process

Theorem 7.10. *Let $(\Delta_n : n \in \mathbb{N})$ be a sequence, such that $\Delta_n \downarrow 0$ and for every n let $0 = t_0^n < t_1^n < \ldots$ with $t_i^n = i\Delta_n$ be a discretization of the time axis. For every n let $\{Y_i^n, i = 1, 2, \cdots\}$ be a sequence of random variables, where Y_i^n are defined as the number of jumps of a Poisson process $(N_t, t \geq 0)$ during the time interval $(t_{i-1}^n, t_i^n]$ and denote $p_i^n(k)$ the probability that $Y_i^n = k$, for $k = 0, 1, \cdots, K$ and $p_i^n(K+1) = 1 - \sum_{k=0}^K p_i^n(k)$. Define $k^n(t) = \sup\{i : t_i^n \leq t\}$ and*

$$M_t^n = \sum_{i=1}^{k^n(t)} Y_i^n.$$

Then M_t^n converges in distribution to N_t.

Proof. Fix $t \geq 0$. It is enough to prove that $\Psi_{M_t^n}(\mathrm{u})$ converges to $\Psi_{N_t}(u)$ for all u. We have $k_n(t) = \left\lfloor \frac{t}{\Delta_n} \right\rfloor$. The characteristic function of Y_i^n is given by

$$\Psi_{Y_i^n}(u) = \mathrm{E}[\exp(\mathrm{i}uY_i^n)] = \sum_{k=0}^{K+1} e^{\mathrm{i}uk} p_i^n(k) = \sum_{k=0}^{K} (e^{\mathrm{i}uk} - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^k}{k!} + e^{\mathrm{i}u(K+1)}.$$

Since the random variables Y_i^n are independent, the characteristic function of their sum is equal to product of characteristic functions of the sumands. Hence

$$\begin{aligned}
\Psi_{M_t^n}(u) &= \prod_{i=1}^{k^n(t)} \Psi_{Y_i^n} = \left\{ \sum_{k=0}^{K} (e^{\mathrm{i}uk} - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^k}{k!} + e^{\mathrm{i}u(K+1)} \right\}^{k^n(t)} \approx \\
&\approx \left\{ \sum_{k=0}^{K} (e^{\mathrm{i}uk} - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^k}{k!} + e^{\mathrm{i}u(K+1)} \right\}^{\frac{t}{\Delta_n}} = \\
&= \exp\left\{ \frac{t}{\Delta_n} \ln \left[\sum_{k=0}^{K} (e^{\mathrm{i}uk} - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^k}{k!} + e^{\mathrm{i}u(K+1)} \right] \right\} = \\
&= \exp\left\{ \frac{t}{\Delta_n} \ln \left[(1 - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} + e^{\mathrm{i}u(K+1)} + \sum_{k=1}^{K} (e^{\mathrm{i}uk} - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^k}{k!} \right] \right\} \\
&= \exp(*).
\end{aligned}$$

Assume that $(1 - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} + e^{\mathrm{i}u(K+1)} \neq 0$. Then

$$\begin{aligned}
(*) &= \frac{t}{\Delta_n} \ln \left[(1 - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} + e^{\mathrm{i}u(K+1)} \right] + \\
&\quad + \frac{t}{\Delta_n} \ln \left[1 + \sum_{k=1}^{K} \frac{(e^{\mathrm{i}uk} - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^k}{k!}}{1 - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} + e^{\mathrm{i}u(K+1)}} \right] \approx \\
&\approx \frac{t}{\Delta_n} \ln(e^{-\lambda\Delta_n}) + \frac{t}{\Delta_n} \ln \left[e^{\mathrm{i}u(K+1)} (e^{\lambda\Delta_n} - 1) + 1 \right] + \\
&\quad + \frac{t}{\Delta_n} \frac{(e^{\mathrm{i}u} - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} \lambda\Delta_n}{(1 - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} + e^{\mathrm{i}u(K+1)}} + \frac{t}{\Delta_n} \sum_{k=2}^{K} \frac{(e^{\mathrm{i}uk} - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} \frac{(\lambda\Delta_n)^k}{k!}}{1 - e^{\mathrm{i}u(K+1)}) e^{-\lambda\Delta_n} + e^{\mathrm{i}u(K+1)}}.
\end{aligned}$$

The first term is equal to $-\lambda t$. The second term, with $\Delta_n \searrow 0$, can be approximated as

$$\frac{t}{\Delta_n} \ln \left[e^{\mathrm{i}u(K+1)} (e^{\lambda\Delta_n} - 1) + 1 \right] \approx \frac{t}{\Delta_n} e^{\mathrm{i}u(K+1)} (e^{\lambda\Delta_n} - 1) \approx \frac{t}{\Delta_n} e^{\mathrm{i}u(K+1)} \lambda\Delta_n = \lambda t e^{\mathrm{i}u(K+1)}.$$

The third term with $\Delta_n \searrow 0$ converges to

$$\lambda t (e^{\mathrm{i}u} - e^{\mathrm{i}u(K+1)}),$$

and the last term converges to zero with $\Delta_n \searrow 0$.

Hence

$$\Psi_{M_t^n}(u) \xrightarrow{\Delta_n \to 0} \exp\{\lambda t (e^{\mathrm{i}u} - 1)\},$$

and this is exactly the characteristic function of the Poisson process N_t with parameter λ. □

Compound Poisson process lattice approximation Consider a compound Poisson process $(L_t : t \geq 0)$ associated with the jump intensity $\lambda > 0$ and the distribution of the jumps σ and a density function f. Its moment generating function is given by

$$\Psi_{L_t}(u) = \exp\left\{ \lambda t (\hat{\sigma}(u) - 1) \right\}. \tag{7.5}$$

The approximation of L_t will be done in 2 steps. First we find the discretized approximation of the distribution σ. Then we apply this discretization and find a lattice approximation of the compound Poisson process L_t. The first step is done in the following two Propositions.

Proposition 7.11. *Let X be a real-valued random variable with values in $\mathbb{R}$ with distribution σ and corresponding distribution function F. Let $\varepsilon \in (0,1)$ be arbitrary. Define*

$$x_L = \sup\{x : \mathbb{P}(X < x) < \varepsilon\} \quad \text{and} \quad x_U = \inf\{x : \mathbb{P}(X > x) < \varepsilon\}.$$

Then a random variable X^ε with distribution function F^ε such that

$$F^\varepsilon(x) = \begin{cases} 0 & \text{for } x < x_L, \\ F(x) & \text{for } x_L \leq x < x_U, \\ 1 & \text{for } x \geq x_U, \end{cases}$$

converges in distribution to X as $\varepsilon \to 0$.

Proof. It's enough to proof that for every continuous and bounded function g we have

$$\lim_{\varepsilon \downarrow 0} \mathrm{E}[g(X^\varepsilon)] = \mathrm{E}[g(X)].$$

This follows easily since $x_L \to -\infty$ and $x_U \to \infty$ as $\varepsilon \downarrow 0$. □

Theorem 7.12. *Let X be a real-valued random variable with values in a bounded interval D with distribution σ and corresponding distribution function F. For every $n \in \mathbb{N}$ we make a partition of D into n subdomains,*

$$\inf D = x_0^n < x_1^n < \cdots < x_n^n = \sup D,$$

such that $\sup_j |x_j^n - x_{j-1}^n| \downarrow 0$ as $n \to \infty$. We define a new variable

$$X^n = x_j^n \quad \text{with probability } p_j^n = F(x_j^n) - F_-(x_{j-1}^n), \text{ for } j = 1, \cdots n.$$

Then $X^n \xrightarrow{\mathrm{d}} X$ as $n \to \infty$.

Proof. Consider an arbitrary function g that is continuous and bounded. We show that $\mathrm{E}[g(X^n)] \to \mathrm{E}[g(X)]$ as $n \to \infty$. We have

$$\mathrm{E}[g(X^n)] = \sum_{j=1}^{n} g(x_j^n)(F(x_j^n) - F_-(x_{j-1}^n)) \tag{7.6}$$

Here, (7.6) is the Riemann–Stieltjes sum that converges with $n \to \infty$ to the integral (7.7)

$$\int_D g(x)\mathrm{d}F(x) = \mathrm{E}[g(X)]. \tag{7.7}$$

We conclude that X^n converges in distribution to X. □

Once we have the approximation of the distribution of jumps, we may continue with the approximation of the compound Poisson process. We make a partition of the time axis and define a new variable, that express if the process jumps in the time interval corresponding to the partition and if so, what is the size of the jump. Here we assume for simplicity of the construction that only one jump can occur during the time interval. This assumption is correct when the length of the time interval goes to zero, however more precise results could be obtained by considering more than one jump. The size of the jump is approximated by the previous Theorem 7.12. The precise result is as follows.

Proposition 7.13. *Let $(P_t : t \geq 0)$ be a compound Poisson process associated with the intensity of jumps $\lambda > 0$ and the distribution of jumps σ. Let $(\Delta_n : n \in \mathbb{N})$ be a sequence such that $\Delta_n \downarrow 0$ and for every n let $0 = t_0^n < t_1^n < \cdots$ with $t_i^n = i\Delta_n$ be a discretization of the time axis. Consider an approximation of the jump size distribution σ as in the Theorem 7.12 (if the distribution has unbounded domain, we first use the Proposition 7.11) and define a new variable Y_i^n such that*

$$Y_i^n = \begin{cases} 0 & \text{with probability } p^n = \exp(-\lambda\Delta_n) \\ x_j^n & \text{with probability } (1-p^n)p_j^n, \quad j = 1, \cdots, n. \end{cases} \tag{7.8}$$

Then P_t^n defined as

$$P_t^n = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} Y_i^n$$

converges in distribution to P_t.

Proof. We compute the characteristic function of P_t^n and show that it converges to characteristic function of the compound Poisson process P_t. For Δ_n in the neigborhood of 0 we may approximate the characteristic function of P_t^n as follows.

$$\begin{aligned}
\Psi_{P_t^n}(u) &= \prod_{i=1}^{\lfloor t/\Delta_n \rfloor} \Psi_{Y_i^n}(u) = \prod_{i=1}^{\lfloor t/\Delta_n \rfloor} \left[p^n + \sum_{j=1}^{n} e^{ux_j^n}(1-p^n)p_j^n \right] = \\
&= \Big[p^n + (1-p^n)\psi_{X^n}(u) \Big]^{\lfloor t/\Delta_n \rfloor} \approx \Big[p^n + (1-p^n)\psi_{X^n}(u) \Big]^{t/\Delta_n} = \\
&= \exp\left\{ \frac{t}{\Delta_n} \log\left[e^{-\lambda\Delta_n}(1 + \psi_{X^n}(u)(e^{\lambda\Delta_n} - 1)) \right] \right\} = \\
&= \exp\left\{ -\lambda t + \frac{t}{\Delta_n} \log\left[1 + \psi_{X^n}(u)(e^{\lambda\Delta_n} - 1) \right] \right\} \approx \\
&\approx \exp\left\{ -\lambda t + \frac{t}{\Delta_n}\psi_{X^n}(u)(e^{\lambda\Delta_n} - 1) \right\} \approx \exp\left\{ -\lambda t + \frac{t}{\Delta_n}\psi_{X^n}(u)\lambda\Delta_n \right\} = \\
&= \exp\left\{ \lambda t(\psi_{X^n}(u) - 1) \right\} \stackrel{n\to\infty}{\to} \exp\{\lambda t(\psi_X(u) - 1)\}.
\end{aligned}$$

Hence P_t^n converges in distribution to the compound Poisson process P_t. □

Inverse Gaussian process lattice approximation Any inverse Gaussian process is a subordinator and a pure jump process, hence its characteristic function is given by

$$\Psi_{IG}(z) = \exp\left(\int_0^\infty (e^{izx} - 1)u_{IG}(x) \right).$$

If we compare this to a characteristic function of a compound Poisson process (2.10)

$$\Psi_{CPP}(z) = \exp\left(\lambda \int_0^\infty (e^{izx} - 1)\sigma(dx) \right)$$

we see a very similar structure. However, inverse Gaussian process itself is not a compound Poisson process as the total mass of its Lévy measure is infinite. On the other hand if we truncate the Lévy density near zero, the total mass becomes finite, and thus we obtain a compound Poisson process. We know from the Proposition 2.8 that every Lévy process is a limit of a sequence of compound Poisson processes, we can thus approximate an IG process by such a sequence. The precise result is as follows.

Proposition 7.14. *Let X be an $IG(\delta,\gamma)$ process with Lévy density $u_{IG}(x)$. Define a new process X^{ε} containing the jumps of X smaller than ε, with characteristic function*

$$\Psi_{X_t^{\varepsilon}}(z) = \exp\left(t\int_0^{\varepsilon}(e^{\mathrm{i}zx}-1)u_{IG}(x)\mathrm{d}x\right).$$

Then $(X - X^{\varepsilon}) \xrightarrow{d} X$ as $\varepsilon \to 0$.

Proof. Follows easily from the fact that $\Psi_{X_t^{\varepsilon}}(z) \xrightarrow{\varepsilon\to 0} 1$. □

Proposition 7.15. *Define a new process $Y = (X - X^{\varepsilon})$. Then Y is a compound Poisson process with jump intensity*

$$\lambda = \int_0^{\infty} u_{IG}(x)\mathbf{1}_{\{|x|>\varepsilon\}}\mathrm{d}x$$

and distribution of jumps

$$\sigma(x) = \frac{1}{\lambda}u_{IG}(x)\mathbf{1}_{\{|x|>\varepsilon\}}.$$

Proof. The characteristic function of Y_1 can be written as

$$\begin{aligned}\Psi_{Y_1}(z) &= \exp\left(\int_{\varepsilon}^{\infty}(e^{\mathrm{i}zx}-1)u_{IG}(x)\mathrm{d}x\right) = \\ &= \exp\left(\int_0^{\infty}(e^{\mathrm{i}zx}-1)u_{IG}(x)\mathbf{1}_{\{|x|>\varepsilon\}}\mathrm{d}x\right).\end{aligned}$$

We have to show that $u_{IG}(x)\mathbf{1}_{|x|>\varepsilon}$ has a finite mass, then Y will be a compound Poisson process.

$$\begin{aligned}\left|\int_0^{\infty} u_{IG}(x)\mathbf{1}_{\{|x|>\varepsilon\}}\mathrm{d}x\right| &= \int_{\varepsilon}^{\infty}\frac{1}{\sqrt{2\pi}}\delta x^{-3/2}\exp\left(-\frac{\gamma^2}{2}x\right)\mathrm{d}x \le \\ &\le \int_{\varepsilon}^{\infty}\frac{1}{\sqrt{2\pi}}\delta x^{-3/2}\mathrm{d}x < \infty.\end{aligned}$$

□

The compound Poisson process $(Y_t : t \ge 0)$ can be approximated by the method described in the previous paragraph. We apply the Proposition 7.13 to approximate the process Y given the intensity λ and jump distribution σ defined in the Proposition 7.15.

7.2.2 Approximation of the IG–OU type process process

Let $a > 0$, $\sigma = 1$ and consider an IG–OU type process defined by (4.3), i.e.

$$X_t = e^{-at}X_0 + \int_0^t e^{-a(t-s)}\mathrm{d}L_s.$$

From the Proposition 4.23 we know, that the stochastic integral

$$I_t = \int_0^t e^{-a(t-s)}\mathrm{d}L_s$$

can be represented as a sum of an inverse Gaussian random variable and a compound Poisson random variable, $I_t = I_t^{(1)} + I_t^{(2)}$, where $I_t^{(1)}$ is $\mathrm{IG}\Big(\delta\left(1-e^{-\frac{1}{2}at}\right),\gamma\Big)$ random variable and $I_t^{(2)}$

is a compound Poisson process with intensity $\delta\left(1-e^{-\frac{1}{2}at}\right)$ and with jumps having common density function

$$f_t(x)=\begin{cases}\frac{\gamma^{-1}}{\sqrt{2\pi}}x^{-3/2}\left(e^{\frac{1}{2}at}-1\right)^{-1}\left(e^{-\frac{1}{2}\gamma^2x}-e^{-\frac{1}{2}\gamma^2xe^{at}}\right), & x>0\\ 0, & \text{otherwise.}\end{cases}\tag{7.9}$$

We know from the Remark 4.7 that an OU–type process is a temporally homogenous Markov process. Hence, for a time interval of length Δ and a starting point $X_t=x$, we can write

$$X_{t+\Delta}=e^{-a\Delta}x+I_\Delta=e^{-a\Delta}x+I_\Delta^{(1)}+I_\Delta^{(2)}\tag{7.10}$$

Consider a sequence $\Delta_n\downarrow 0$ and a partition Δ of the time interval $[0,t]$, $0=t_0^n<t_1^n<\cdots<t_{k_n}^n=t$ with $t_j^n=j\Delta_n$ (we can always choose Δ_n such that $k_n\Delta_n=t$). Using the methods described in the previous subsections, we approximate the processes $I_t^{(1)}$ and $I_t^{(2)}$ by Y_t^n and P_t^n using the partition Δ, and define a sequence of processes $(I_t^n:t\geq 0)_{n\in\mathbb{N}}=(Y_t^n+P_t^n:t\geq 0)_{n\in\mathbb{N}}$. From (7.10) follows that given the value $X_{t_j^n}=x$ the next value of the approximated IG–OU type process is determined as

$$X^n_{t^n_{j+1}}=e^{-a\Delta_n}x+I^n_{\Delta_n}.$$

The term $I^n_{\Delta_n}$ is a discrete random variable not depending on t whose distribution is given in the following result.

Theorem 7.16. *In the previous setting, the difference $I^n_{\Delta_n}$ is a discrete random variable. Choose $\varepsilon\in(0,1)$ and consider the following notation.*

(1) The process P_t is a compound Poisson process associated with the jump intensity

$$\lambda=\delta\gamma(1-e^{-\frac{1}{2}at})$$

and the distribution of jumps f given by (7.9). Denote the corresponding distribution function by F and consider $I=[0,F^{-1}(1-\varepsilon))$. Consider a partition of I into n subdomains

$$0=r_0^n<r_1^n<\cdots<r_n^n=\sup I$$

such that the mesh is going to 0 with n going to infinity. Define

$$R_i^n=r_i^n\qquad\text{with probability } p_i^n=F(r_i^n)-F(r_{i-1}^n),\qquad\text{for } i=1,\dots,n.$$

Denote

$$p^n=e^{-\lambda\Delta_n}.$$

(2) The process Y_t is a compound Poisson process associated with the jump intensity

$$\kappa=\int_\varepsilon^\infty u_{IG}(x)\mathrm{d}x$$

and the distribution of jumps

$$g(x)=\frac{1}{\kappa}u_{IG}(x)\mathbf{1}_{\{x>\varepsilon\}},\qquad x>0.$$

Denote the corresponding distribution function by G and consider $J = [\varepsilon, G^{-1}(1-\varepsilon))$. Consider a partition of J into n subdomains

$$\varepsilon = x_0^n < x_1^n < \cdots < x_n^n = \sup J$$

such that the mesh is going to 0 with n going to infinity. Define

$$X_i^n = x_i^n \qquad \text{with probability } q_i^n = G(x_i^n) - G(x_{i-1}^n), \qquad \text{for } i = 1, \ldots, n.$$

Denote

$$q^n = e^{-\kappa \Delta_n}.$$

Then

$$I_{\Delta_n}^n = Y_{\Delta_n}^n + P_{\Delta_n}^n$$

is a random variable, taking the following values

$$\begin{array}{llll} i_{00}^n = 0 & \text{with probability} & pi_{00}^n = p^n \times q^n & \\ i_{i0}^n = r_i^n & \text{with probability} & pi_{i0}^n = (1-p^n)p_i^n \times q^n, & i = 1, \ldots, n \\ i_{0k}^n = x_i^n & \text{with probability} & pi_{0k}^n = p^n \times (1-q^n)q_k^n, & k = 1, \ldots, n \\ i_{ik}^n = r_i^n + x_k^n & \text{with probability} & pi_{ik}^n = (1-p^n)p_i^n \times (1-q^n)q_k^n, & i,k = 1, \ldots, n. \end{array}$$

Example 7.17. Here we give an example of the approximation of the IG–OU type process using the previous Theorem. For different values of parameters a, δ, γ, ε and Δ_n we approximated the distribution of jumps of the process Y_t^n and P_t^n by 10 values and for every j we decided by generating a uniform random variable if the processes Y_t^n and P_t^n jumps and if so what is the size of the jumps based on the approximation. Results of such approximation are given in the Figure 7.7. By comparing this Figure with Figure 7.1 we see a the same structure – moves up entirely by jumps and then tails off exponentially. The Figure 7.8 shows the empirical autocorrelogram of the approximated process compared to the theoretical autocorrelation function given by

$$r(s) = \operatorname{corr}(X_t, X_{t+s}) = e^{-as}.$$

Finally the Figure 7.9 compares the empirical density function of the approximated IG–OU type process with its theoretical density function. From the figures we conclude that the approximation of the IG–OU type process is quite good, but the distribution is not fitted perfectly. This can be improved by not neglecting the positive probability of the occurence of more than one jump in the time period Δ_n.

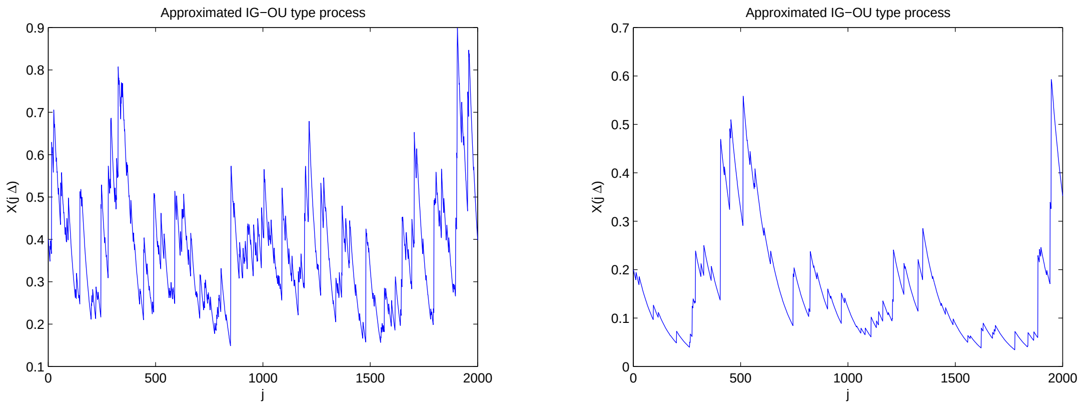

Figure 7.7: Approximated IG–OU type process $X(j\Delta)$ against j. Left: $a = 0.02, \delta = 2, \gamma = 5$ and $\Delta = 1$. Right: $a = 0.01, \delta = 1, \gamma = 5$ and $\Delta = 1$.

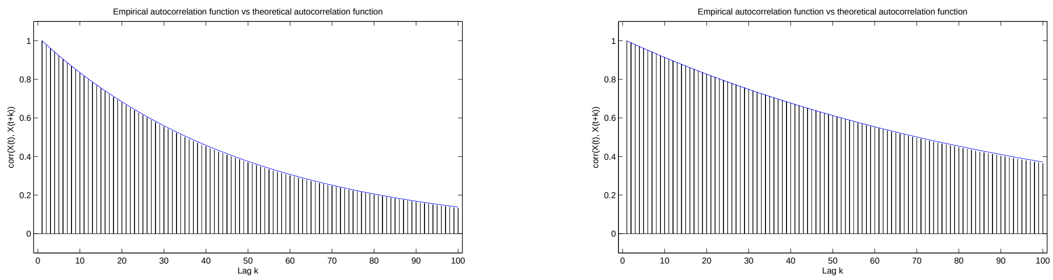

Figure 7.8: Empirical autocorrelation function of the approximated data $X(j\Delta)$, $j = 0, \ldots, 200000$, vs. the theoretical autocorrelation function. Left: $a = 0.02, \delta = 2, \gamma = 5$ and $\Delta = 1$. Right: $a = 0.01, \delta = 1, \gamma = 5$ and $\Delta = 1$.

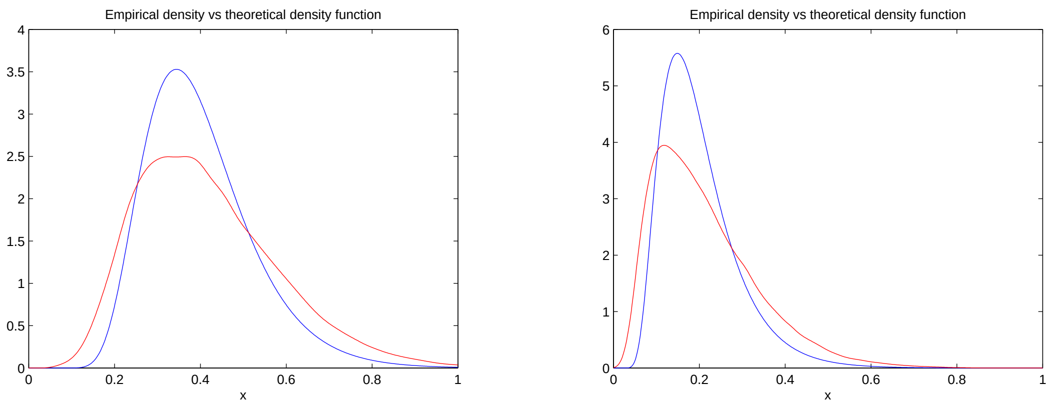

Figure 7.9: Empirical (red) and theoretical (blue) density function of the approximated IG–OU type process. Left: $a = 0.02, \delta = 2, \gamma = 5$ and $\Delta = 1$. Right: $a = 0.01, \delta = 1, \gamma = 5$ and $\Delta = 1$.

7.2.3 Approximation of the short rate IG–OU type process

We assume the generalized Vašíček model of the short rate, where the short rate is the IG(δ,γ)–OU type process with drift. We assume without loss of generality that $\bar{\sigma}=1$. The short rate is given by

$$r_t = m_t + \int_0^t e^{-a(t-s)}\mathrm{d}L_s,$$

where

$$m_t = f(0,t) - \frac{\delta}{\gamma}\frac{(1-e^{-at})}{\sqrt{1+2a^{-1}(1-e^{-at})/\gamma^2}}.$$

We constructed the approximation of the stochastic term I_t in the previous subsection. The approximation lattice r^n of r then takes values

$$r^n_{t^n_j} = m_{t^n_j} + I^n_{t^n_j}.$$

From a node $r^n_{t^n_j}$ at time t^n_j it branches to nodes

$$r^n_{t^n_{j+1}} = e^{-a\Delta_n}r^n_{t^n_j} + am_{t^n_{j+1}} - ae^{-a\Delta_n}m_{t^n_j} + i^n_{ik}, \quad i,k=0,\dots,n$$

with probabilities $pi^n_{i,k}$ given in the Theorem 7.16.

This method gives us an explicitely given lattice approximation of the short rate process that could be used for pricing interest rate derivatives in the following way. Consider a contingent claim with a payoff $C_T = C(r_T)$. Then the value at time t of this contingent claim is given by the risk neutral pricing formula as

$$C_t = \mathrm{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s\mathrm{d}s\right)C_T|\mathcal{F}_t\right].$$

Once we have a lattice approximation r^n of the short rate process then we can approximate this value by backward induction, i.e.

$$C^n_{t^n_j} = \sum_{i,k=0}^{n} pi^n_{ik}\exp\left(e^{-a\Delta_n}r^n_{t^n_j} + am_{t^n_{j+1}} - ae^{-a\Delta_n}m_{t^n_j} + i^n_{ik}\right)C^n_{t^n_{j+1}}.$$

This way we can go back to time 0 and get the price of the contingent claim at any time.

Although this method looks simple, it needs a very high order branching and the lattice is not recombining. This leads to very quick growth of computer capacity needed for the computation. Making the lattice recombining is quite difficult since the jumps of the compound poisson processes Y^n_t and P^n_t comes from a very different distributions, the first having only small jumps and the other having only big jumps.

Conclusion

In the present thesis we generalized the HJM model of term structure of interest rates by considering general Lévy process as a driver. We developed a theory for the construction of arbitrage-free short rate model for general Lévy driver and general volatility structure. By specifying the volatility structure we obtained Ornstein–Uhlenbeck type short rate process that posseses the basic property of interest rates – mean reversion. We focused our attention to two specific models. The IG–OU type model where the short rate process is assumed to be an IG–OU type process with drift implies positive interest rates (under some mild conditions on the parameters of the IG distribution), is mean reverting and easy to simulate. The drawback of this model is that it allows only positive jumps. We constructed a lattice approximation of such process. The other model we studied was OU–NIG model where the short rate follows an OU type process with drift and the driver is a NIG process. This process is also mean reverting, however it can become negative with a positive probability due to the negative jumps of the driving Lévy process. It is also easy to simulate and a method of parameters estimation was shown.

Problem of estimation of the model parameters arises with the IG–OU model. There has been an extensive study of estimation methods for the OU type processes, see for example Valdivieso et al. (2007) for the maximum likelihood estimation or Spiliopoulos (2008) for the method of moments method. Maximum likelihood method relies on the knowledge of the density function that is not known in the IG–OU model of the short rate. Generalized method of moments requires a stationary process that is also not the case of the IG–OU short rate.

Other problem to think about is incorporating the negative jumps. We have seen that the OU–NIG model fits really well the financial data as shows the calibration on yield curves data in the Chapter 6. However the negative interest rates are not acceptable. A possibility would be to study the generalization of the CIR model that in the Brownian case leads to positive short rate, or truncate the negative jumps of the NIG driver.

Bibliography

Anděl, J. (2005). *Základy matematické statistiky (In Czech).* Matfyzpress, Prague, Czech Republic.

Applebaum, D. (2004). *Lévy processes and stochastic calculus.* Cambridge University Press, Cambridge, United Kingdom.

Barndorff-Nielsen, O. E. (1998). Process of normal inverse Gaussian type. *Finance and Stochastics*, 2:41–68.

Barndorff-Nielsen, O. E. and Halgreen, C. (1977). Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, 38:309–311.

Barndorff-Nielsen, O. E., Jensen, J. L., and Sørensen, M. (1998). Some stationary processes in discrete and continuous time. *Advances in Applied Probability*, 30:998–1007.

Barndorff-Nielsen, O. E. and Shephard, N. (2000). Modelling by Lévy processes for financial econometrics. In O. E. Barndorff-Nielsen, T. M. and Resnick, S., editors, *Lévy processes - Theory and Applications*, pages 283–318. Birkhäuser, Boston, USA.

Bauer, H. (1996). *Probability theory.* De Gruyter Studies in Mathematics, Berlin, Germany.

Bertoin, J. (1996). *Lévy processes.* Cambridge University Press, Cambridge, United Kingdom.

Cont, R. and Tankov, P. (2003). *Financial modelling with jump processes.* Chapmann & Hall/CRC, London, United Kingdom.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53:385–407.

Dothan, L. (1978). On the term structure of interest rates. *Journal of Financial Economics*, 6:59–69.

Eberlein, E. (2000). Application of generalized hyperbolic Lévy motions to finance. In O. E. Barndorff-Nielsen, T. M. and Resnick, S., editors, *Lévy processes - Theory and Applications*, pages 319–336. Birkhäuser, Boston, USA.

Eberlein, E. and Kluge, W. (2006). Exact pricing formulae for caps and swaptions in a Lévy term structure model. *The Journal of Computational Finance*, 9(2):99–125.

Eberlein, E. and Kluge, W. (2007). Calibration of Lévy term structure models. In Fu, M. C., Jarrow, R. A., Yen, J.-Y., and Elliott, R. J., editors, *Advances in Mathematical Finance*, pages 147–172. Birkhäuser, Boston, USA.

Eberlein, E. and Raible, S. (1999). Term structure models driven by general Lévy processes. *Mathematical Finance*, 9(1):31–53.

Grosswald, E. (1976). The Student t-distribution of any degree of freedom is infinitely divisible. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, 36:103–109.

Heath, D., Jarrow, R. A., and Morton, A. (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica*, 60:77–105.

Ho, T. S. Y. and Lee, S.-B. (1986). Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41:1011–1029.

Hull, J. and White, A. (1990). Pricing interest-rate derivative securities. *The Review of Financial studies*, 3(4):573–592.

Jacod, J. and Shiryaev, A. N. (2003). *Limit theorems for stochastic processes*. Springer-Verlag, Berlin, Germany.

Jurek, Z. J. and Vervaat, W. (1983). An integral representation for selfdocomposable Banach space random variables. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, 62:247–262.

Lebedev, N. N. (1965). *Special functions and their applications*. Prentice-Hall, inc., Englewood Cliffs, N.J.

Lukacs, E. (1970). *Characteristic functions*. Griffin, London, England.

Madan, D. and Senete, E. (1990). The variance gamma model for share market returns. *Journal of Business*, 63:511–524.

Michael, J. R., Schucany, W. R., and Haas, R. W. (1976). Generating random variates using transformations with multiple roots. *American Statistician*, 30:88–90.

Protter, P. E. (2004). *Stochastic integration and differential equations*. Springer-Verlag, Berlin, Germany.

Raible, S. (2000). *Lévy processes in finance: theory, numerics and empirical facts*. PhD thesis, University Freiburg.

Rudin, W. (1970). *Real and Complex analysis*. McGraw-Hill, London, United Kingdom.

Rydberg, T. H. (1997). The normal inverse Gaussian Lévy processes: simulation and approximation. *Stochastic models*, 14:887–910.

Sato, K. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge University Press, Cambridge, United Kingdom.

Schmidt, W. M. (1997). On a general class of one-factor models for the term structure of interest rates. *Finance and Stochastics*, 1:3–24.

Schoutens, W. (2003). *Lévy processes in finance: Pricing financial derivatives*, chapter 5. John Wiley & Sons, Ltd., West Sussex, England.

Spiliopoulos, K. (2008). Method of moments estimation of Ornstein-Uhlenbeck processes driven by general Lévy process. `http://arxiv.org/abs/0807.2832`.

Vašíček, O. (1977). An equilibrium characterisation of term structure. *Journal of Financial Economics*, 5:177–188.

Valdivieso, L., Schoutens, W., and Tuerlinckx, F. (2007). Maximum likelyhood estimation in processes of Ornstein-Uhlenbeck type. *Statistical Inference for Stochastic Processes*, 12:1–19.

Zhang, S. and Zhang, X. (2008). Exact simulation of IG–OU process. *Methodology and Computing in Applied Probability*, 10:337–355.

Appendix A

Bessel functions

In this Appendix we follow (Lebedev, 1965, Chapter 5).

Definition A.1. A *Bessel function* is an arbitrary solution of the second order linear differential equation

$$u'' + \frac{1}{z}u' + \left(1 - \frac{\nu^2}{z^2}\right)u = 0, \tag{A.1}$$

where z is a complex variable and ν is a parameter which can take arbitrary real or complex values. The equation (A.1) is calles *Bessel's equation of order ν.*

A general solution to a Bessel equation (A.1) has a form of a linear combination of two arbitrary solutions, i.e.

$$u = C_1 u_1(z) + C_2 u_2(z).$$

Bessel functions of the first and second kind One of the solutions to this Bessel equation is a function J_ν called the *Bessel function of the first kind of order ν*, given by the following formula

$$J_\nu(z) = \sum_{k=0}^{\infty} \frac{(-1)^k (z/2)^{\nu+2k}}{\Gamma(k+1)\Gamma(k+\nu+1)}, \qquad |z| < \infty,\ |\arg z| < \pi. \tag{A.2}$$

To obtain a general solution to the Bessel equation, we introduce a function Y_ν called the *Bessel function of the second kind of order ν* which is defined by the formula

$$Y_\nu(z) = \frac{J_\nu(z)\cos\nu\pi - J_{-\nu}(z)}{\sin\nu\pi} \tag{A.3}$$

for arbitrary z belonging to the plane cut along the segment $[-\infty, 0]$. This function is a solution of (A.1) and is lineary independent with J_ν. Thus the expression

$$u = C_1 J_\nu(z) + C_2 Y_\nu(z)$$

is a general Bessel function.

Bessel functions of the third kind The Bessel functions of the third kind are denoted by $H_\nu^{(1)}(z)$ and $H_\nu^{(2)}(z)$ and are defined in terms of the Bessel functions of the first and second kind by the formulas

$$H_\nu^{(1)}(z) = J_\nu(z) + \mathrm{i}Y_\nu(z), \qquad H_\nu^{(2)}(z) = J_\nu(z) - \mathrm{i}Y_\nu(z), \tag{A.4}$$

where ν is arbitrary and z is any point of the plane cut along the segment $[-\infty, 0]$.

The *modified Bessel function of the third kind* or *Macdonald's function* is defined by the following formula

$$K_\nu(z) = \frac{\pi\mathrm{i}}{2} e^{\nu\pi\mathrm{i}/2} H_\nu^{(1)}(ze^{\pi\mathrm{i}/2}), \qquad -\pi < \arg z < \frac{\pi}{2}. \tag{A.5}$$

The function $K_\nu(z)$ satisfy the following recurrence relations, see (Lebedev, 1965, Chapter 5, Section 5.8) equation (5.7.9), (5.7.10).

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[z^\nu K_\nu(z)\right] = -z^\nu K_{\nu-1}(z), \qquad \frac{\mathrm{d}}{\mathrm{d}z}\left[z^{-\nu} K_\nu(z)\right] = -z^{-\nu} K_{\nu+1}(z). \tag{A.6}$$

The following formula holds

$$K_{-\nu}(z) = K_\nu(z). \tag{A.7}$$

Integral representation of K_ν The modified Bessel function has several integral represetations, see (Lebedev, 1965, Chapter 5, Section 5.10) equations (5.10.23) – (5.10.25).

$$K_\nu(z) = \int_0^\infty e^{-z\cosh u} \cosh \nu u \,\mathrm{d}u, \qquad |\arg z| < \frac{\pi}{2}, \qquad \nu \in \mathbb{C}, \tag{A.8}$$

$$K_\nu(z) = \frac{\sqrt{\pi}}{\Gamma(\nu + \frac{1}{2})}\left(\frac{z}{2}\right)^\nu \int_1^\infty e^{-zt}(t^2-1)^{\nu-1/2}\mathrm{d}t, \qquad |\arg z| < \frac{\pi}{2}, \qquad Re(\nu) > -\frac{1}{2}, \tag{A.9}$$

and finally

$$K_\nu(z) = \frac{1}{2}\int_0^\infty y^{\nu-1} \exp\left\{-\frac{1}{2}z(y+y^{-1})\right\}\mathrm{d}y, \qquad |\arg z| < \frac{\pi}{4}, \qquad \nu \in \mathbb{C}. \tag{A.10}$$

Series representation of K_ν Combining the series representation of J_ν given by (A.2). If $\nu \notin \mathbb{Z}^+$ then we can use the definition of $Y_\nu(z)$ together with the series representation of J_ν given by (A.2). If $\nu = n \in \mathbb{Z}^+$ the fraction (A.3) is to be understood in the "l'Hôpital" sense; here Y_ν has the following series representation (see (Lebedev, 1965, Chapter 5, Section 5.5) equation (5.5.3)).

$$\begin{aligned} Y_n(z) = {} & \frac{2}{\pi}J_n(z)\log\frac{z}{2} - \frac{1}{\pi}\sum_{k=0}^{n-1}\frac{(n-k-1)!}{k!}\left(\frac{z}{2}\right)^{2k-n} \\ & - \frac{1}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k(z/2)^{n+2k}}{k!(n+k)!}\left[\gamma(k+1)+\gamma(k+n+1)\right], \qquad |\arg z| < \pi,\ n \in \mathbb{N}, \end{aligned} \tag{A.11}$$

where

$$\gamma(k) = \frac{\Gamma'(k)}{\Gamma(k)}.$$

Using these representation we obtain the series representation for $K_\nu(z)$. The zero order terms yield the behaviour of K_ν near 0

$$\begin{aligned} K_\nu(x) &\approx \frac{2^{\nu-1}\Gamma(\nu)}{x^\nu}, \qquad x \to 0, \\ K_0(x) &\approx \log\frac{2}{x}, \qquad x \to 0. \end{aligned}$$

Asymptotic representation of K_ν Asymptotic formulas for the modified Bessel function K_ν is derived from the intgral representation (A.9) and is given by

$$K_\nu(z) = \left(\frac{\pi}{2z}\right)^{1/2} e^{-z}\left[\sum_{k=0}^{n}(\nu,k)(2z)^{-k} + O(|z|)^{-n-1}\right], \qquad |\arg z| \le \pi - \delta, \tag{A.12}$$

where δ is arbitrary small and (ν,k) is given by

$$(\nu,k) = \frac{(-1)^k}{k!}\left(\frac{1}{2}-\nu\right)_k\left(\frac{1}{2}+\nu\right)_k,$$

where $(\alpha)_k$ is the Pochhammer symbol of any complex α, defined by

$$(\alpha)_0 = 1, \quad \text{and} \quad (\alpha)_{k+1} = (\alpha+k)(\alpha)_k,$$

for any integer $k \ge 0$; see (Lebedev, 1965, Chapter 5, Section 5.11), equation (5.11.9). From the leading term of (A.12) we determine the asymptotic behaviour as $x \to \infty$ in the form

$$K_\nu(x) \approx \sqrt{\frac{\pi}{2x}}e^{-x}, \qquad x \to \infty.$$

Specialization of K_ν We specialize to the case when $\nu = n + \frac{1}{2}$, $n \in \mathbb{Z}^+$. Assume that $\mathrm{Re}(z) > 0$. Then, using the integral representation (A.9)

$$K_\nu(z) = \frac{\sqrt{\pi}}{\Gamma(n+1)}\left(\frac{z}{2}\right)^{n+1/2}\int_1^\infty e^{-zt}(t^2-1)^n \mathrm{d}t$$

and using the binomial formula we get

$$K_\nu(z) = \sum_{k=0}^{n}\binom{n}{k}(-1)^{n-k}\frac{\sqrt{\pi}}{n!}\left(\frac{z}{w}\right)^{n+1/2}\int_1^\infty e^{-zt}t^{2k}\mathrm{d}t,$$

where the integrals are elementary functions given as linear combinations of functions of type $e^{-z}/z^l, l \in \mathbb{Z}^+$. In the case when $\nu = \frac{1}{2}$ this yields

$$K_{1/2}(z) = \sqrt{\pi}\sqrt{\frac{z}{2}}\int_1^\infty e^{-zt}\mathrm{d}t = \sqrt{\frac{\pi}{2z}}e^{-z}.$$